INSIGHTS INTO THE INTRODUCTION AND DISTRIBUTION OF INVASIVE JAPANESE CLIMBING FERN (*LYGODIUM JAPONICUM*) THROUGH WHOLE CHLOROPLAST SEQUENCING

A Thesis by

Morgan Lynn Cox

Bachelor of Science, Wichita State University, 2019

Submitted to the Department of Biological Sciences
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Master of Science

May 2022

INSIGHTS INTO THE INTRODUCTION AND DISTRIBUTION OF INVASIVE JAPANESE CLIMBING FERN (*LYGODIUM JAPONICUM*) THROUGH WHOLE CHLOROPLAST SEQUENCING

The following faculty members have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Science, with a major in Biological Sciences.

James Beck Committee Chair

Bin Shuai, Committee Member

Crystal Dozier, Committee Member

Michael McKain, Committee Member

DEDICATION

This thesis work is dedicated to my husband, Benjamin. Thank you for being my go-to study buddy since undergrad (even when you had no idea what I was studying). I would not be here today without your continued support, encouragement and sacrifice. Here's to the next chapter in our lives.

ACKNOWLEDGMENTS

First and foremost, I want to thank Dr. James Beck for his invaluable advice, continued support, and patience during my time in the Beck Lab. Your knowledge and experience have encouraged me during academic research and daily life. I wouldn't be in this position without the numerous white board diagrams, and editing comments. The time spent in your lab was ferntastic.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Bin Shuai, Dr. Crystal Dozier and Dr. Michael McKain for your insightful comments and questions.

I would particularly like to extend my gratitude to Ethan Altergott. I would have been an even *more* stressed out grad student without your diligent work over the summer performing library protocols.

This research also would not have been possible without the countless collectors, staff, volunteers, and donors that make herbaria possible. Special thanks to Tiana Franklin Rehman for being so gracious, and helpful during our visits to the BRIT herbarium.

I also extend my gratitude to; fellow lab mates Stacy Holt and Barnabas Hawkinson, for our time spent working together; Dr. Kiley Hicks, for making a TA's job as painless as possible given the ever-changing Covid guidelines; Dr. Mary Liz Jameson, for her constant encouragement and positive attitude; and to my fellow Bio grad students, for our time spent together (in person and virtual) over the past two years.

Lastly, I would like to acknowledge the countless friends and family who supported me over the years. It means the world that you always listen to my "cool science facts."

ABSTRACT

Japanese Climbing Fern (*Lygodium japonicum*) is a vine native to the open forests of eastern Asia that has become an invasive species in the United States. Herbarium records suggest Florida or North Carolina are the initial site of introduction during the early 1900's, but the fern is now established in much of the southeast including Alabama, Arkansas, Georgia, Louisiana, Mississippi, South Carolina and Texas. We aimed to ask three questions regarding the introduction of *L. japonicum*: (1) Was there a single Japanese Climbing Fern introduction or multiple introductions? (2) What is the distribution of haplotypes in the United States? and (3) What are the source population(s) from the native range in Asia? To answer these questions, we sequenced the chloroplast genome from 74 *L. japonicum* herbarium specimens representing 24 native and 50 invasive range populations. Seventeen haplotypes were found in the native range compared to three in the invasive range. Our results indicate *L. japonicum* has low genotype diversity in the invasive range relative to the native range. Even with low genotype diversity, this data suggests at least three introductions of *L. japonicum*. However, the native source population(s) of invasive *L. japonicum* remains unknown. These findings add to our understanding of invasive species introductions, and may have implications for biological control.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Background and significance

Invasive species can result in both ecological and economical detriments. While impacts vary, invasive plants alone have been shown to decrease local plant species diversity and alter nutrient cycling (Vilà et al. 2011). Meta analyses suggest that by the time shifts in community structure and changes in nutrient cycling are detected, major impacts have occurred to native plant species and adverse effects are likely to continue (Vilà et al. 2011). Over the past 500 years, analyses show that non-native species are the second most common threat associated with extinction for five major taxa- plants, amphibians, reptiles, birds and mammals (Bellard et al. 2016). The detrimental effects of non-native species were exceeded only by human mediated habitat destruction. In addition to their negative effects on biodiversity, invasive plant species also negatively impact agriculture. Complications with farming and forestry can result in monetary losses and a decline in human health. The estimated 50,000 non-native species in the United States result in losses totaling $137 billion per year (Pimentel et al. 2000).

There are multiple hypotheses regarding what makes an introduced species successful at invasion. These characteristics ultimately stem from three factors: (1) patterns of genetic diversity in the species' native range, (2) the number and location of introductions, and (3) the number of founding individuals per introduction (Dlugosch and Parker 2008; Shirk et al. 2014). Regarding founding events, these often include only a fraction of the genetic variants that occur in the source population (Dlugosch and Parker 2008). Therefore, it is predicted that multiple introductions, especially from different source populations, could result in increased genetic variation in the introduced range (Genton et al. 2005; Dlugosch and Parker 2008). These effects,

however, may be species specific. For example, if the native range has low geographic genetic differentiation, then it is possible that the even a single introduced population contains most of the genetic variation of the native range. This means that a single introduction event of a nonnative species could be detrimental to the native community (Dlugosch and Parker 2008; Simon et al. 2011). Alternatively, if a single or small number of introductions leads to reduced variation, the result may be a lack of variation in resistance loci to biological control or environmental factors such as drought or temperature extremes (Dlugosch and Parker 2008; Kurose et al. 2020). This genetic bottlenecking (a significant reduction in genetic diversity due to a reduction in population size) may only be a temporary constraint. Many species are introduced through trafficking and commerce transportation- vectors that are repeatedly used. This leads to a high probability that there will be multiple introductions. Indeed, only one example of a potential single introduction of a plant invasive species could be found in the literature (Saltonstall 2002), strongly suggesting that multiple introductions are more common. As the number of introductions increases, and admixture occurs between different lineages- novel genetic variants or heterozygous loci could increase mean fitness of the invasive (Colautti and Lau 2015). Species often become well established in new environments long before detection, resulting in a lack of knowledge regarding their history and dynamics. Uncovering the origin, number of introductions, and genetic diversity of invasive species is important for testing hypotheses regarding invasion, spread, and control. In this study we aim to establish the number of introductions, source area, and geographic distribution of genotypes in an invasive fern species.

The fern genus *Lygodium* is relatively small; comprising 40 species, including one species native to North America [*Lygodium palmatum* (Bernh.) Sw.] and many from the Eastern Hemisphere (Lott et al. 2003). Japanese climbing fern [*Lygodium japonicum* (Thunb.) Sw.] is a

vine native to the open forests of eastern and northeastern Asia, the East Indies, and northern Australia (Wyatt and Wyatt 2013). *Lygodium japonicum* has become an invasive species in the United States, expanding rapidly in the 20th century. Herbarium records suggest Florida or North Carolina are the initial site of introduction in the early 1900's, but the fern is now established in much of the southeast including Alabama, Arkansas, Georgia, Louisiana, Mississippi, South Carolina and Texas (Fig. 1) (Hutchinson and Langeland 2010; Wyatt and Wyatt 2013). The fern is found growing in damp areas such as forests, along waterways, and in disturbed roadside ditches. Japanese climbing fern has twining fronds of indeterminant growth- due to the lack of terminal structures, petioles continue to elongate indefinitely. In fact, records show that petioles have reached 90 feet in length (Minogue et al. 2009). This growth pattern allows for the plant to create a dense canopy that effectively shades out underlying native vegetation. As such, *L. japonicum* is a threat to native plants in pine and hardwood forest communities.

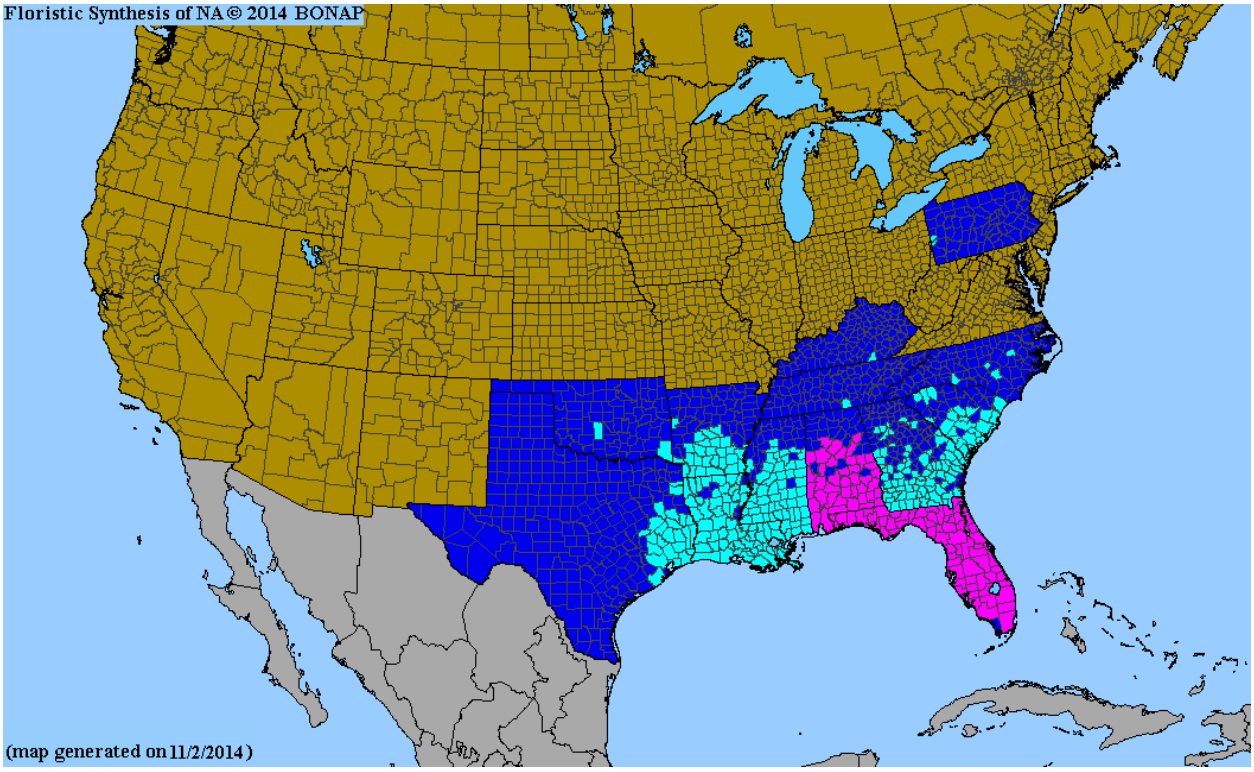

Figure 1. A Biota of North America Program (BONAP) range map of *Lygodium japonicum* in the United States. Dark blue indicates states were *L. japonicum* is present, while light blue indicates the presence at the county level. Magenta denotes states where the species is considered noxious (Kartesz 2015).

Many reproductive life history characteristics of non-native plants have been proposed to facilitate greater competitive ability and eventual invasion. Such characteristics include self-fertilization, rapid growth to reproductive age, high and/or continuous seed (or spore) production, and adaptions for dispersal (Lott et al. 2003). Unfortunately, *L. japonicum* possesses many of these characteristics. There are three modes of reproduction in ferns: intragametophytic selfing (selfing of an individual gametophyte), intergametophytic selfing (reproduction between two gametophytes from the same sporophyte) and intergametophytic crossing (reproduction between two gametophytes from different sporophytes) (Soltis and Soltis 1992; Lott et al. 2003; Haufler et al. 2016). Lott et al. (2003) found that over 90% of isolated *L. japonicum* gametophytes produced successful sporophytes via intragametophytic selfing. This means that

even a single spore could suffice to found a new Japanese climbing fern population. Due to this reproductive potential and its ability to reach above tree canopies, long distance wind dispersal and colonization of *L. japonicum* may be achieved by the successful establishment of single spores.

Ongoing efforts to control Japanese climbing fern lack specificity. Most studies focus on the closely related Old World climbing fern [*Lygodium microphyllum* (Cav.) R. Br.], with few studies focusing on the Japanese climbing fern. Insect species of interest for potential biological control include the lygodium gall mite (*Floracarus perrepae* Knihinicki & Boczek), lygodium saw fly (*Neostrombocerus* sp.), flea beetles (*Manobia* sp.), and stem borers (Minogue et al. 2009). The caterpillar of *Neomusotima fuscolinealis* Yoshiyasu is a natural pest of *L. japonicum* in its native range of Japan. However, due to potential environmental safety precautions, the insect has yet to be tested as a biological control in the United States (Pemberton 1998). Importantly, the lack of information regarding *L. japonicum* genetic variation in the U.S. limits current biological control efforts. Knowledge of genotypic diversity, origin(s) and genotype distribution is important for controlling invasive plants, especially if there is preliminary evidence that control agents have differential success across multiple genotypes (Gaskin et al. 2005). Indeed, a study looking at multiple *Lygodium* species found that successful biological control of one species was not effective on other species within the genus due to genetic diversity (Madeira et al. 2008). Additionally, current prescribed fires and burning pose little detriment to the species and may provide a fuel ladder to canopy trees. Because the Japanese climbing fern is rooted in the ground, the vegetative growth upward provides a mode (ladder) for the fire to climb from the ground to the canopy. Their underground rhizomes also allow for survival and fast regrowth of vegetation following a burn. The rhizome is also thought to play a key role in frost

survival because mild winters do not necessarily kill below-ground portions of the plant (Minogue et al. 2009). As such, the use of herbicides may be required in addition to prescribed fires to slow the spread of *L. japonicum* (Leichty et al. 2011). Due to the current lack of biological and mechanical controls, it is likely that *L. japonicum* will continue to spread and displace native vegetation throughout the southeast United States. The only distribution or range limitation may be rainfall and consistently low winter temperatures (Wyatt and Wyatt 2013).

1.2 Aims

We will pursue three aims regarding the introduction of *L. japonicum*. Our first aim is to determine if there was a single Japanese climbing fern introduction or multiple introductions. If it has been introduced multiple times, we will document the distribution of the alternative haplotypes throughout the southeastern United States. Our final aim is to determine the source population(s) from the native range in Asia. These three aims will all be accomplished by sequencing whole chloroplast genomes from herbarium specimens throughout the distribution of *L. japonicum* in the United States and east Asia. Why the focus on herbarium specimens? Sampling tissue from herbarium specimens allows us to quickly assemble a geographically rigorous set of samples. There simply isn't enough time to field-collect samples from both *L. japonicum*'s introduced and (especially) native range in the course of a M.S. thesis.

This raises the question, though, can DNA sufficient for genomic techniques be routinely obtained from herbarium specimens? The vast majority of specimens were not collected with the intention of DNA extraction, and specimens may have therefore been subject to damaging treatments such as collecting in alcohol, drying at high heat, or pest control practices such as freezing. This question has prompted numerous researchers to study the relationship between extracted DNA quality and various aspects of herbarium specimens, with the goal of identifying factors that predict which specimens will produce higher quality DNA. One such factor is leaf

color (greenness). Erkens et al. (2008) extracted DNA from over 200 herbarium specimens representing multiple genera and specimen ages from 3 to 430 years. They found no significant relationship between DNA extraction success and a specimen's leaf color. In addition, they found leaf color to be affected by many factors including how the specimen was collected and taxon. Other factors may also include chemical composition of leaf tissue, adaption to drought stress, or other biological causes. Kates et al. (2021) also studied "greenness" as a predictor of DNA sequencing success for nearly 8,0000 herbarium specimens. Their sampling included 23 families, although ca. 60% of specimens were from Fabaceae. DNA extraction yield were quantified with PicoGreen fluorescence and sequencing success was measured as the number of on-target reads divided by the total number of reads (proportion-on-target reads recovered). Proportion-on-target reads recovered was calculated for each sample by mapping trimmed reads to 229 reference nuclear sequences (Kates et al. 2021). Using multivariate linear mixed regression models, they evaluated the relationships between multiple predictors (greenness, specimen age, herbarium and absolute latitude) and DNA yield and proportion-on-target reads. They concluded that greenness was the best predictor for DNA yield, however color alone was not a clear predictor for sequencing success.

Another herbarium specimen characteristic thought to be a predictor for DNA extraction is specimen age. The age of specimens may hinder DNA extraction success due to the natural degradation of DNA over time (Staats et al. 2011). The study of ca. 8,000 specimens by Kates et al. (2021) spanned a specimen age range of 2 to 182 years (mean 40 years). They concluded that specimen age had no significant effect on DNA yield. However, specimen age in relationship to "greenness" was indicative of proportion-on-target reads recovered. In younger samples greenness was associated with lower proportion-on-target reads recovered, and as sample age

increased greener samples were associated with higher proportion-on-target reads recovered (Kates et al. 2021). Bakker et al. (2015) extracted and sequenced DNA from 93 specimens (73 herbarium specimens) with ages up to 146 years old. They also found no correlation between specimen age and DNA yield. However, fresh and silica dried specimens had slightly more total number of reads compared to herbarium specimens. In summary, multiple studies have concluded that although various aspects of herbarium specimens can affect sequencing success, these studies suggest that this presents no strong barrier to sequencing from herbarium-derived DNA (Erkens et al. 2008; Staats et al. 2011; Kates et al. 2021). These findings reveal that even historic herbarium specimens from the Linnaean era can be useful in future genomic studies (Andreasen et al. 2009).

Our specific goal is obtaining whole chloroplast genome sequences, and fortunately previous studies strongly suggest that these can be obtained from herbarium material. Regardless of tissue type, the chloroplast genome is relatively easy to sequence due to its abundance in the cell. Whole chloroplast genomes are therefore easily obtained by relatively low coverage shotgun sequencing of genomic libraries, so called "genome skimming" (Steele and Pires 2011; Straub et al. 2012). With regards to herbarium specimens, previous research indicates that these specimens do not exhibit a significant difference in the degree of DNA fragmentation between the chloroplast, mitochondrial, and nuclear genomes (Staats et al. 2011). Furthermore, since the production of genomic libraries begins with enzymatically fragmenting the DNA, the already highly fragmented DNAs obtained from herbarium specimens are compatible with genome skimming. Alsos et al. (2020) evaluated the success of genome skimming in a broad range of herbarium specimens from the European Alps- 6,655 specimens representing 161 families. Overall, they were able to assemble the full chloroplast genome for 67% of the samples (Alsos et

al. 2020). Successful genome skimming also doesn't require much material. Zeng et al. (2018) focused on how feasible it was to minimize herbarium specimen destruction and still recover chloroplast genomes. Their study confirmed DNA extracted from small amounts of tissue (1 cm^2) could result in successful chloroplast genome assemblies (Zeng et al. 2018).

Additionally, prior studies have shown chloroplast DNA sequence to be useful for identifying the number of introductions, genetic variation and the geographic distribution of invasive genotypes. Note that with regards to the haploid chloroplast genome, the genotype is often referred to as a haplotype. One study examined 996 bp from two highly variable chloroplast DNA regions to establish haplotypes in the invasive species *Lepidium draba* L. (Gaskin et al. 2005). Forty-one different haplotypes were found in the 684 specimens collected from the native Eurasian range and the introduced US range. Twenty of these haplotypes were found in the U.S., clearly suggesting that multiple introductions have occurred. This study also found that most of the genetic variation in both the native and introduced range was observed within geographical regions, not between regions.

Hufbauer and Sforza (2007) focused on two Eurasian knapweed species that are invasive to North America, *Centaurea diffusa* and *Centaurea stoebe* (Asteraceae). Leaf tissue from 30 native European locations and 20 invasive North American locations were field collected. A total of 873 bp of chloroplast DNA was sequenced from 213 samples. *Centaurea diffusa* exhibited 11 haplotypes, with 9 observed in the native range and only three in the invasive range (Hufbauer and Sforza 2007). These data suggest that native range genetic diversity is higher than that seen in the invasive range, although it does appear that multiple introductions to North America have occurred. Interestingly, only one a haplotype (the most common North American haplotype) was observed in both the native and introduced ranges. In comparison, the *C. stoebe* sample set

contained 11 haplotypes- 10 in the native range, and four in the invasive range. All invasive haplotypes were also found in Bulgaria. Hufbauer and Sforza (2007) suggest that *C. stoebe* invasive range diversity is best explained by multiple introductions from a narrow geographical region in the native range. Taken together, these studies set up several expectations for our study- namely that multiple introductions are to be expected, but that a reduced portion of native range genetic variation will be present in the invasive range (Table 1).

Existing data suggests that chloroplast genome sequencing will be a successful strategy for understanding *L. japonicum* invasion genomics. Importantly, a published chloroplast genome is available (Gao et al. 2013), which will aid in genome assembly. We also expect to find sufficient variation in the chloroplast genome. The studies above all found sufficient variation in one or a few individual chloroplast regions, and we will be sequencing the entire chloroplast genome. Indeed, a previous study has successfully used single chloroplast DNA genes to reconstruct an intra-specific phylogeny for another invasive *Lygodium* species (*L. microphyllum*) (Madeira et al. 2008). This suggests that sufficient variation for identifying multiple origins will be present in the *L. japonicum* chloroplast genome.

Table 1. Comparable studies using chloroplast DNA to investigate U.S. plant invasions.

Taxon	Invasion path	Reference	Invasive haplotypes/ samples sequenced	Invasive haplotype frequency	bp Sequenced	Common haplotype observed in native range?
Lepidium draba	Europe to US	Gaskin 2005	20/341	1 every 17	996 bp	yes
Phragmites australis	Europe to US	Saltonstall 2002	13/257	1 every 20	2,000	yes
Brassica nigra	Eurasia to North America	Oduor 2015	13/90	1 every 7	>1,000 bp	yes
Centaurea diffusa	Eurasia to US	Hufbauer 2008	3/33	1 every 11	873 bp	yes
Centaurea stoebe micranthos	Eurasia to US	Hufbauer 2008	4/36	1 every 9	873 bp	yes
Polysiphonia harveyi	Japan to North America	McIvor 2001	2/3	1 every 1.5	1,245 bp	no
Schinus terebinthifollius	South America to US	Williams 2005	2/354	1 every 177	716 bp	yes
Salvinia molesta	South America to US	Holt unpublished	9 in 21	1 every 2.3	whole cp	no
Lygodium japonicum	Asia to US	this study	3 in 50	1 every 17	whole cp	no

CHAPTER 2
METHODS

2.1 Obtaining samples

All samples were obtained exclusively from herbarium specimens. Our goal was to broadly sample across *L. japonicum*'s native and invasive range. Herbarium samples were obtained through loans and in-person visits. A visit to the Botanical Research Institute of Texas herbarium (BRIT), including the Vanderbilt University (VDB) and University of Louisiana-Monroe (NLU) herbaria, in late 2020 provided much of the invasive-range sampling. Loans from regional herbaria, The University of South Carolina Herbarium (USCH) and the University of Florida Herbarium (FLAS), were used to supplement this initial invasive-range sampling. A loan from the University and Jepson Herbarium of the University of California, Berkeley (UC) provided much of the native range sampling.

Each *L. japoncium* herbarium specimen was examined to confirm species identification. Key characteristics of *L. japonicum* include twining leaves that are twice pinnately compound, with pinnae (leaflet) up to 12 inches wide, and pinnules (subleaflet) up to 3 inches long. Fertile and sterile pinnae are subdimorphic; fertile leaves are usually smaller and have finger like projections around the margins containing sporangia (Munger 2005). In most cases, a single specimen was selected for each county for which material was available, and approximately one half of one leaflet was removed and stored in silica gel. Information was collected from the specimen's herbarium label- collector name, collection date, locality and habitat. Annotation labels were applied to all sheets from which material was removed. Google Earth Pro (Google LLC 2021) combined with online placename searches were used to estimate latitude and longitude coordinates (georeferencing) for all specimen localities. Once all specimens were

georeferenced, the map function in RStudio (RStudio team 2021) was used to visualize the location of all samples.

This map of specimen locations was used to narrow down a list of samples for DNA extraction, and to select specimens with minimum geographic overlap. From the selected samples, ca. 20 mg of tissue was loaded into a 96 well plate. DNA extraction was performed using a standard CTAB protocol modified for 96 well plates (Beck et al. 2012). A Qubit fluorometer (Life Technologies, Eugene, Oregon) was used to establish DNA concentration for all extracts.

2.2 DNA sequencing and alignment

Samples were chosen for genomic library preparation based on geographic disparity and DNA concentration. Library preparations were performed using the NEBNext Ultra II DNA Library Prep Kit for Illumina with the NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1) (NEB, Ipswich, Massachusetts). Library preparation followed the protocol outlined in Saeidi et al. (2018), with 200 ng of input DNA. Samples with low library concentrations were re-amplified (Saeidi et al. 2018) with universal Illumina primers prior to the hybridization reaction. Unenriched libraries were sequenced with 150 bp paired end chemistry on an Illumina (Illumina, San Diego, California) NextSeq 550 at the University of Kansas Genome Sequencing Core. Following de-multiplexing, adapters and low-quality sequence were removed with Trimmomatic (Bolger et al. 2014) implemented on the Kansas State University Beocat high-performance computer. Trimmed plastome sequences were then aligned to a published *Lygodium* chloroplast genome (Gao et al. 2013) using Geneious (Biomatters, Aukland, New Zealand). Sample exhibiting less than 10,000 reads aligned to the *Lygodium* chloroplast genome were removed, as were four samples with highly divergent chloroplast genomes.

Consensus sequences were formed using a threshold of 75% and nucleotides were called as ambiguous if coverage was less than 5. Consensus sequences were aligned using MAFFT (Katoh et al. 2002). Following alignment, all nucleotide positions exhibiting ambiguities, gaps, and identical bases were masked. A TCS (Clement et al. 2002) network was produced from the resulting masked alignment in PopArt (Population Analysis with Reticulate Trees) (Leigh and Bryant 2015). In order to visualize the annotated *L. japonicum* chloroplast genome, the sample exhibiting the fewest ambiguities (ILL71) was aligned to an annotated *L. microphyllum* genome (McCulloch et al. 2018) using MAFFT. Annotations were then transferred to the *L. japonicum* sequence and the sequence was circularized.

CHAPTER 3

RESULTS

3.1 Tissue Sampling

Tissue sampling included 246 *Lygodium* specimens spanning 11 countries, 8 US states and 175 counties (Fig. 2). The collection year of specimens ranged from 1910-2017 (mean = 40.33 years old, ± 25.55 years old). Of the 191 specimens selected for extraction, 189 yielded a measurable DNA concentration. DNA concentration ranged from 0 to 247 ng/µl (mean = 37.64 ng/µl , ± 41.08 ng/µl). DNA was successfully extracted from both the oldest specimen (137 years old), as well as the youngest specimen (3 years old). The relationship between DNA concentration and specimen age is significant (P = 0.037 Fig. 3). However, this relationship is thought to be an artifact from extracting DNA on two separate occasions.

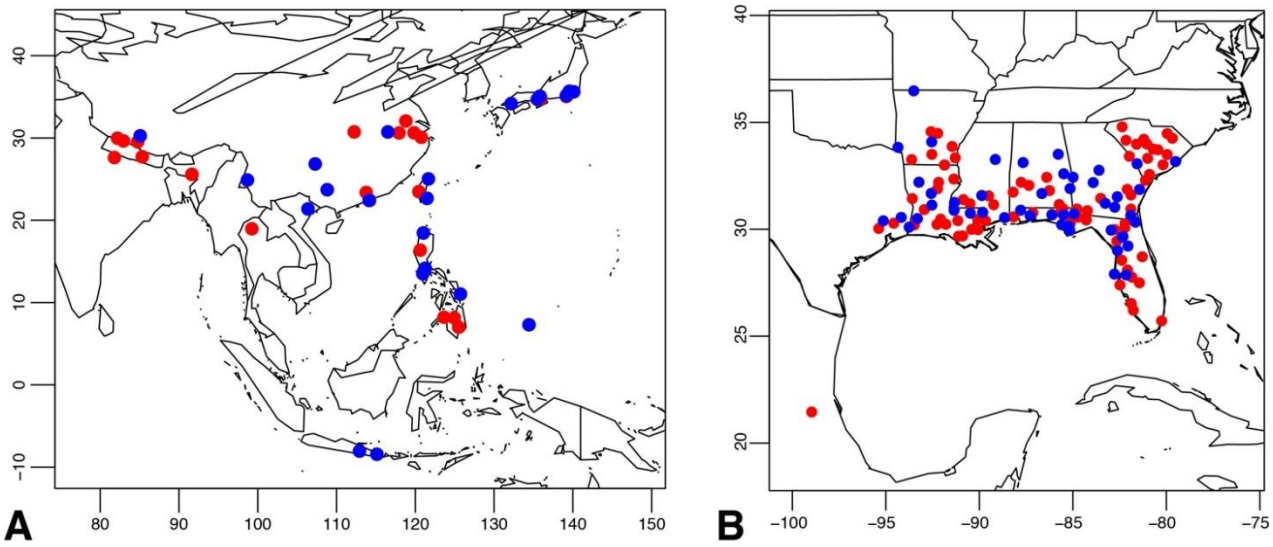

Figure 2. Geographic location of samples for which libraries were made (red) and sequenced (blue). A. Native range. B. Invasive range.

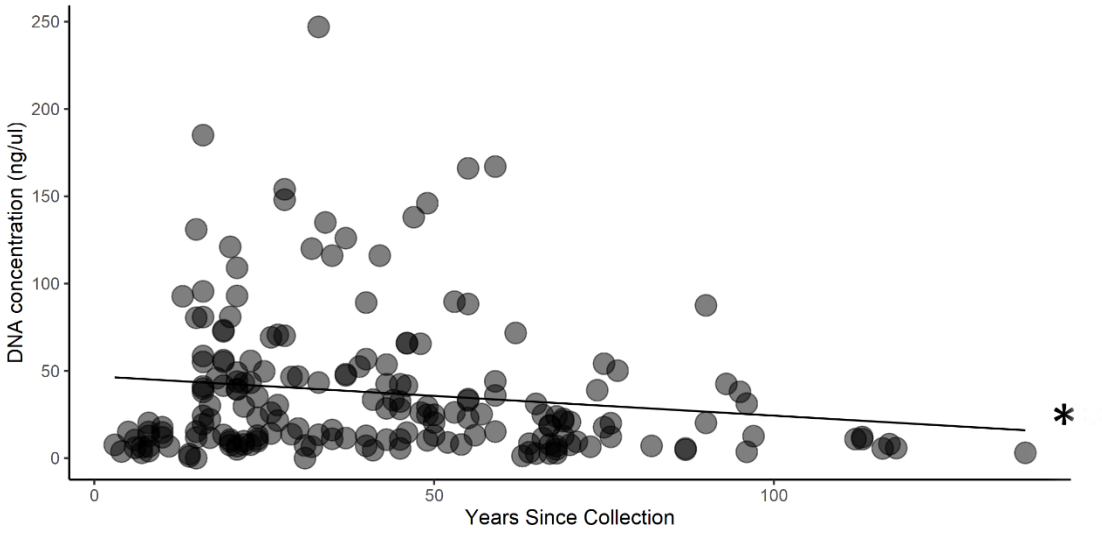

Figure 3. Age had a significant effect on extracted DNA concentration, with older specimens exhibiting lower concentrations (P = 0.037).

After selecting a subset of DNA extractions based on geographic disparity and DNA concentration, 98 samples were chosen for genomic library preparation. Library concentration ranged from 0 - 22.6 ng/µl (mean = 5.84 ng/µl, ± 4.06 ng/µl). Four of the 98 samples failed to yield a measurable DNA concentration via a Qubit fluorometer. However, the relationship between post library DNA concentration and specimen age was not significant (P = 0.485 Fig. 4).

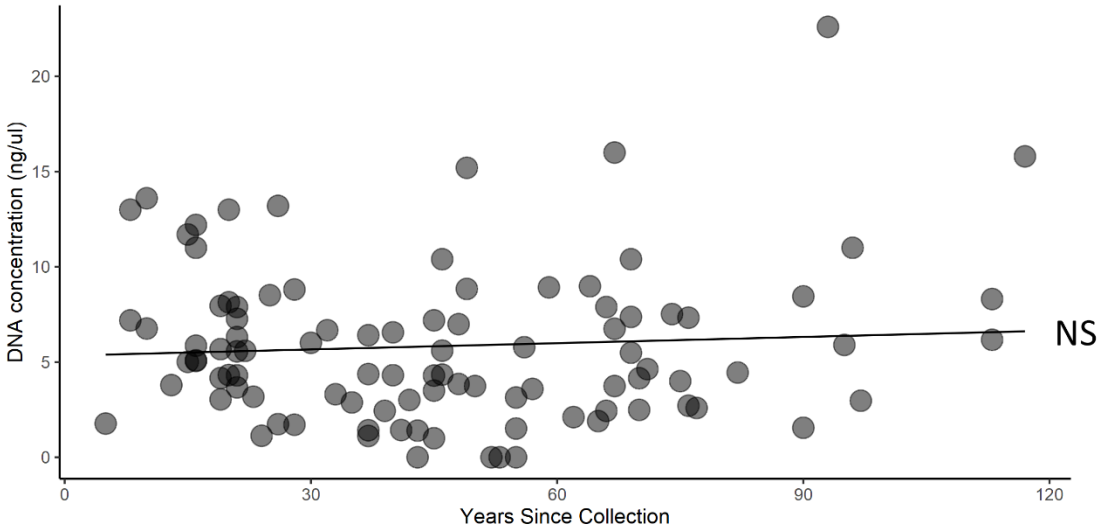

Figure 4. Age did not have a significant effect on library concentration (P = 0.485).

3.2 DNA sequencing

Following library preparation, 87 of the 98 samples were selected for sequencing, and later aligned to the reference chloroplast genome. The number of raw reads ranged from 39,574 - 59,718,342 (mean = 17,099,872 ± 12,443,710). The number of reads mapped to the reference genome ranged from 800 – 806,545 (mean = 174,398 ± 165658). Mean coverage ranged from 0.81 - 679.77 (mean = 129.74 ± 134.85). The total number of ambiguities was highly variable and ranged from 34- 150,406 (mean = 7,726.38 ± 24,957.60). Specimen age had a strong negative effect on sequencing results- including number of raw reads ($P = 2.03e^{-06}$), the number of sequences successfully mapped to the reference genome ($P = 5.03e^{-06}$), and mean coverage ($P = 5.03e^{-06}$) (Figures 5-7). Even with this effect of specimen age, most samples yielded useable chloroplast assemblies, with 85 of 87 samples passing our threshold of >10,000 reads aligned to the *Lygodium* chloroplast genome.

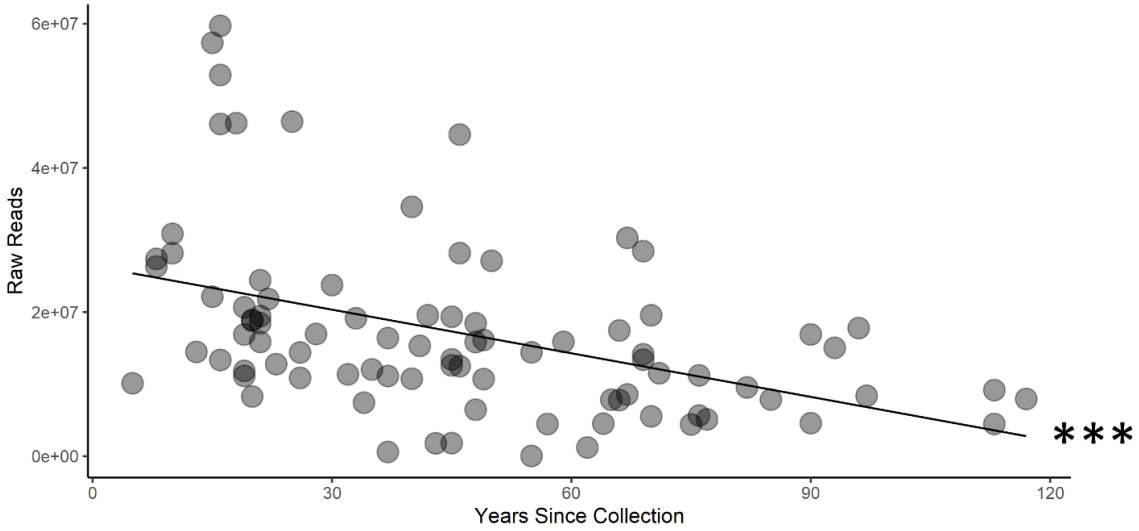

Figure 5. Age significantly affected the number of raw reads obtained, with older specimens resulting in fewer raw reads ($P = 2.03e^{-06}$).

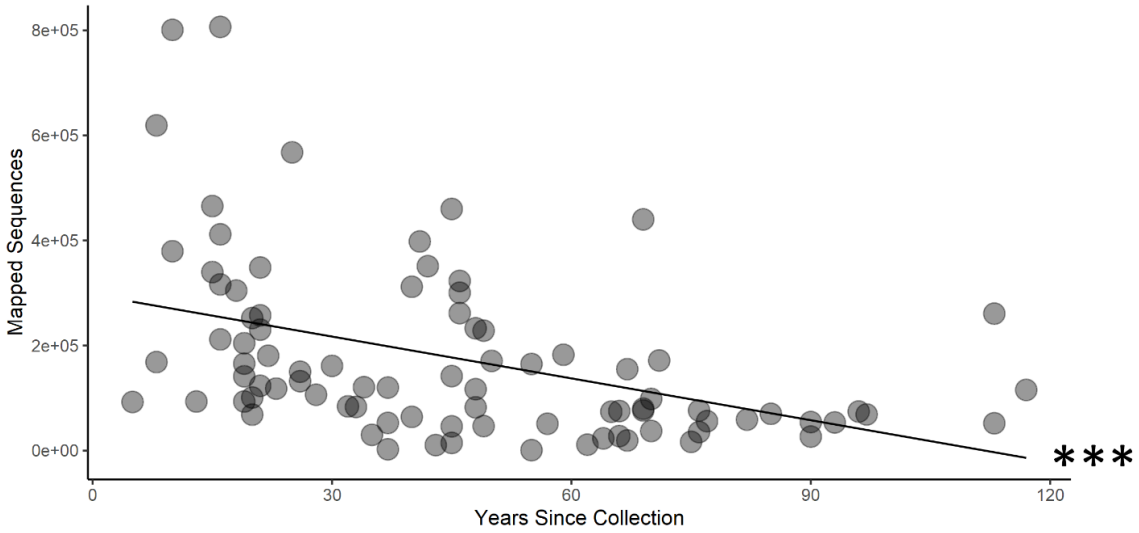

Figure 6. Age significantly affected the number of reads mapped to the reference genome, with older specimens resulting in fewer mapped reads ($P = 5.03e^{-06}$).

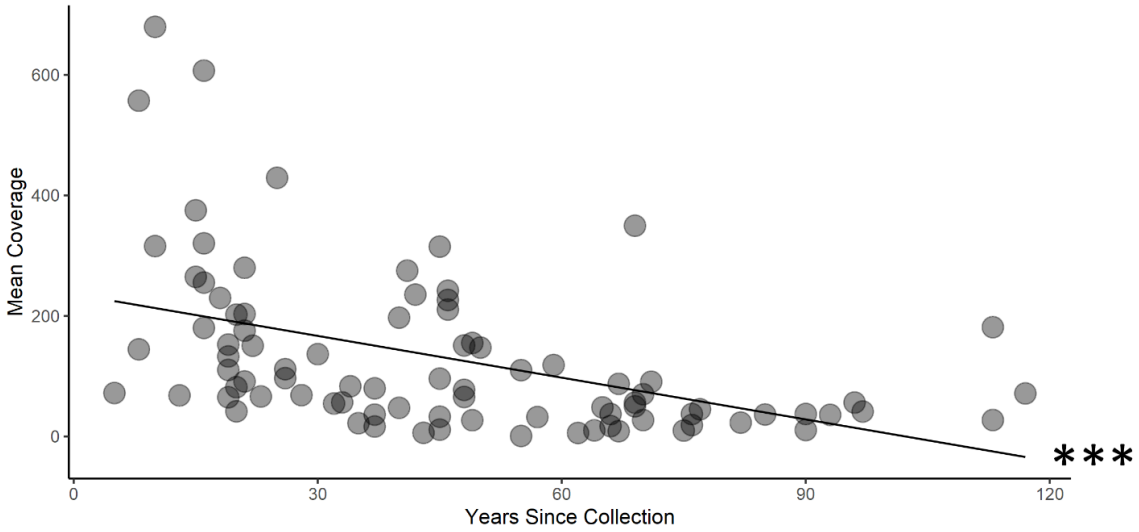

Figure 7. Age significantly affected mean coverage, with older specimens resulting in lower coverage ($P = 5.03e^{-06}$).

Consensus sequences of 74 *L. japonicum* samples with less than 10,000 ambiguities were aligned using MAFFTT (Katoh et al. 2002) in Geneious using default parameters. These consensus sequences included 50 *L. japonicum* samples from the invasive US range and 24 *L. japonicum* samples from the native Asian range. The number of reads mapped ranged from

27,592 - 806,545 (mean = 196,436.10 ± 16,9032.30). Mean coverage ranged from 17.19- 679.77 (mean = 143.37 ± 137.90). The total number of ambiguities - ranged from 34 -679.77 (mean = 763.73 ± 1,383.16). An annotated *L. japonicum* chloroplast genome appears in Fig. 8.

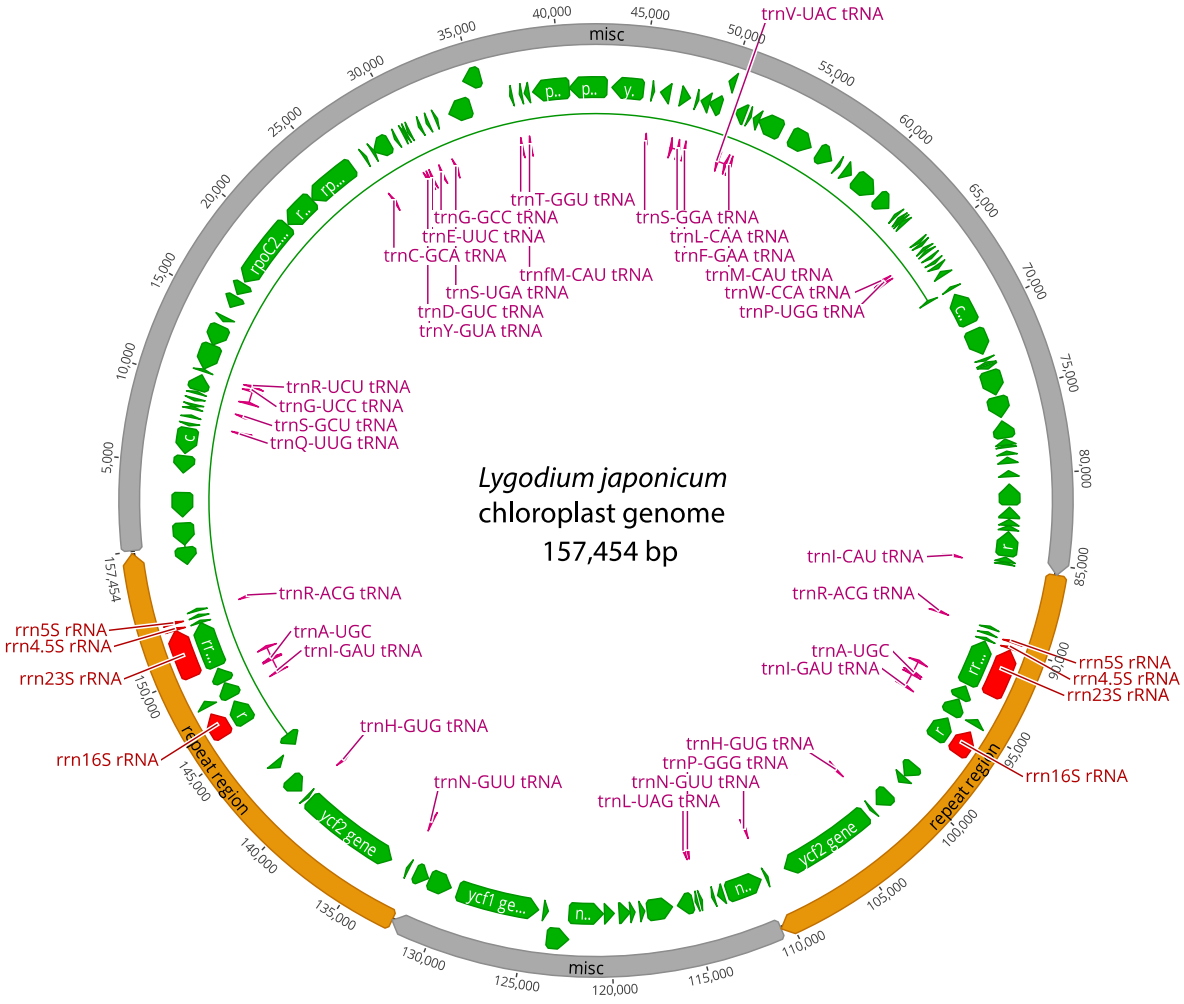

Figure 8. Annotated *Lygodium japonicum* chloroplast genome.

3.3 Haplotype Network

For our network including outgroups, following removal of all ambiguities, gaps, and identical bases from the alignment, 109 single nucleotide polymorphisms (SNPs) remained. This network is shown in Fig. 9. In it, Japanese climbing fern is clearly a genetically distinct unit. For our *L. japonicum*-only network, 35 SNPs remained following the removal of removal of all ambiguities, gaps, and identical bases from the alignment. The TCS network (Fig. 10) exhibited 18 *L. japonicum* haplotypes. Sixteen haplotypes were found in the native range, with three present in the invasive range. A single haplotype dominates the invasion and is found in 47/50 (94%) of US specimens (present in all 8 states sampled). It is also the only haplotype seen in South Carolina, Florida, Alabama, Georgia, Texas and Arkansas, Interestingly, this common haplotype was not observed in the native range. The second invasive haplotype was only found in Mississippi, with the third invasive haplotype present in Louisiana as well as China and the Philippines (Fig. 11).

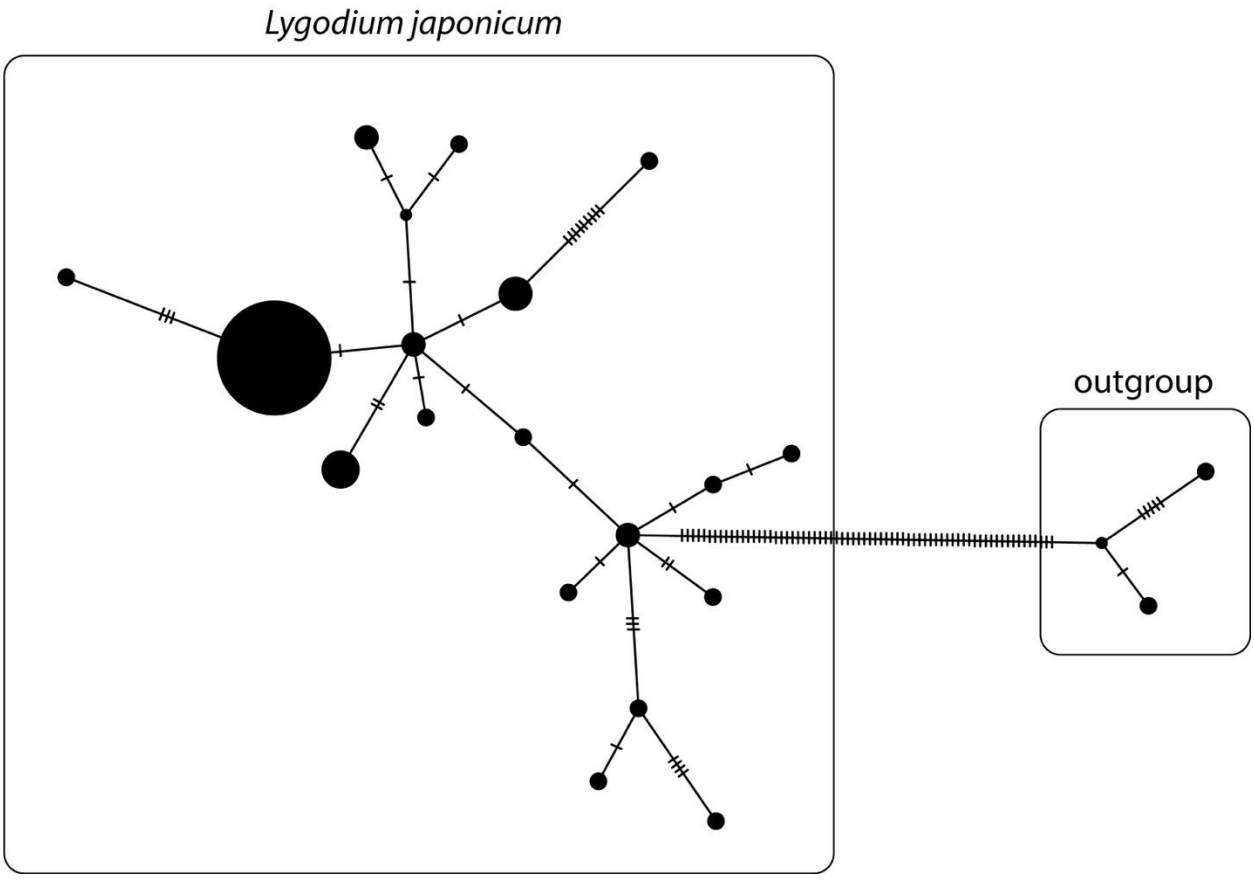

Figure 9. *Lygodium japonicum* haplotype network including two outgroup taxa.

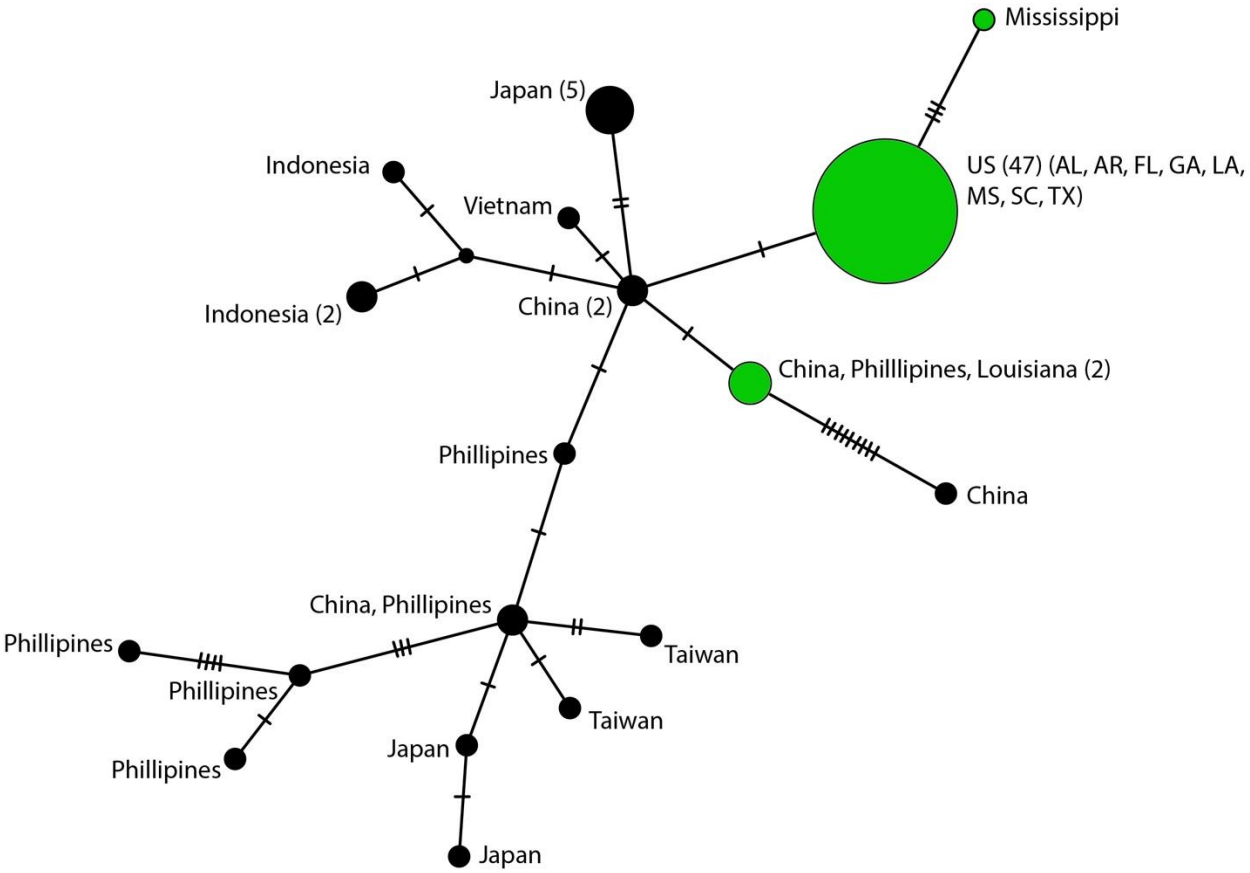

Figure 10. *Lygodium japonicum* haplotype network. Circles represent observed haplotypes and are scaled to frequency. Black circles denote haplotype observed in the native range only, green circles denote haplotypes observed in the invasive range. Tick marks denote nucleotide substitutions.

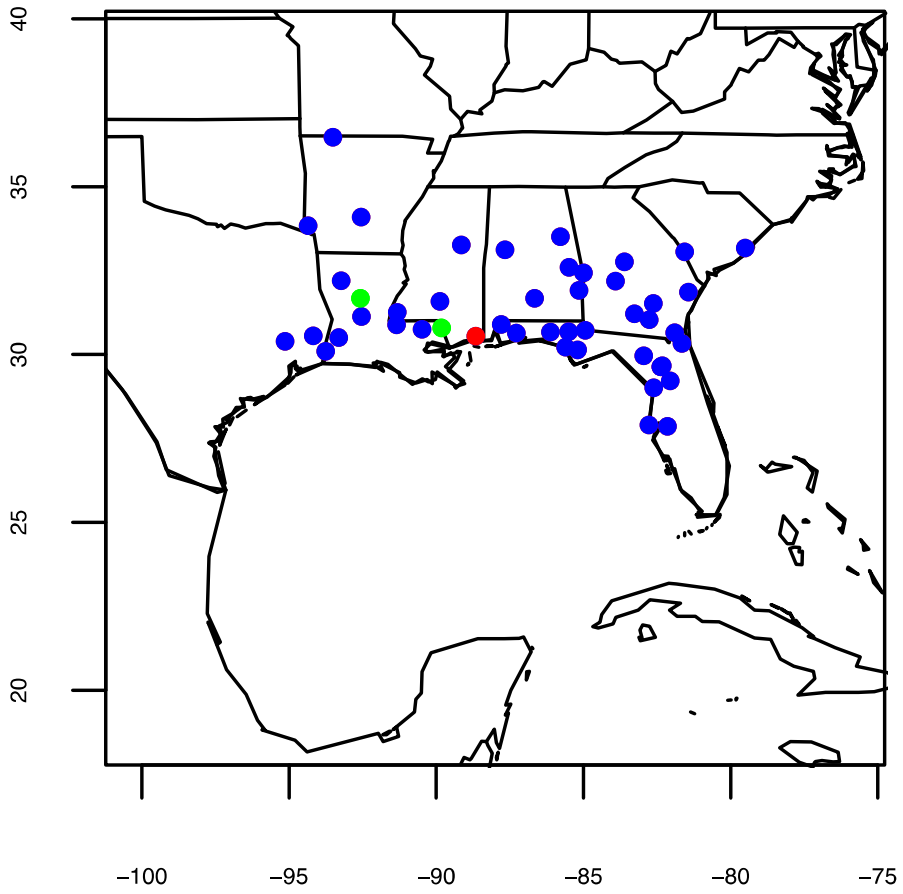

Figure 11. Location of *Lygodium japonicum* invasive haplotypes.

CHAPTER 4
DISCUSSION

4.1 Using herbarium specimens

Multiple studies have concluded that although various aspects of herbarium specimens can affect sequencing success, there is no strong barrier to sequencing from herbarium-derived DNA (Erkens et al. 2008; Staats et al. 2011; Bakker et al. 2015; Kates et al. 2021). Our study also demonstrates the successful use of herbarium specimens for whole chloroplast genome sequencing. We found that age affected a variety of upstream measures of sequencing success, including DNA concentration, the number of raw reads, the number of mapped reads, and mean coverage. However, we were still able to obtain useable chloroplast genome sequences from 85 of 87 sequenced specimens, suggesting that the data loss observed with increasing age is not significant to preclude successful genome assembly.

4.2 Genetic Diversity

We draw three basic conclusions from our genetic results: (1) Japanese climbing fern exhibits a low diversity invasion, (2) there were at least three introductions of *L. japonicum* in the United States, and (3) the native source population(s) of *L. japonicum* remains unknown. It has been suggested that *L. japonicum* was introduced to the United States via trade ships in the early 1900's. Due to the repeated use of such commercial transportation, we expected to observe support for multiple introductions and high genetic diversity within the introduced range. This pattern of multiple introduction and high genetic diversity has been seen in many North American invasive plant species including hoary cress (*Lepidium draba* L.), black mustard (*Brassica nigra* W.D.J. Koch), Eurasian knapweed (*Cenaurea stoebe* L.), and common ragweed (*Ambrosia artmisiifolia* L.) (Table 1). We observed a reduction in genetic diversity in the

invasive (3 haplotypes) vs. the native (16 haplotypes) range (Figure 10). Comparison with other similar studies, albeit limited, suggests that this is a low-diversity invasion. We found 7 studies that examined chloroplast haplotype diversity in a U.S. invasion, and on average an invasive range haplotype was found every 35 samples sequenced (range 1.5-177) (Table 1). In *L. japonicum* we observed a unique haplotype every 17 invasive range samples sequenced, even though we sequenced the entire chloroplast genome (ca. 157,260 bp) instead of the one or two genes (>1500bp) sequenced in these 7 studies. The only other study that examined whole chloroplast genome diversity in a North American invasion is that of *Salvinia molesta* (Holt unpublished data), in which a unique haplotype was observed every 2 samples sequenced. The relative lack of diversity among the *L. japonicum* genomes we sequenced suggests that this is a low diversity invasion.

Our observation of three distinct invasive-range haplotype establishes that multiple introductions have occurred. The only other scenario that would lead to this diversity would be a single introduction comprising multiple haplotypes. We view this as unlikely. Instead, the dominance of the common invasive haplotype suggests that it arrived first and spread throughout the U.S. before the arrival of the two remaining invasive haplotypes. This hypothesis of multiple introductions is consistent with the 8 studies noted above, as all observed multiple invasive range haplotypes (Table 1).

A single haplotype dominated the invasive range, found in 94% of samples in 8 states. The second most common invasive haplotype, present in 2 U.S. samples, was sampled from Louisiana, China and the Philippines. This suggests two possible source populations. Lastly, the least common haplotype was sampled from a single individual from Mississippi, and had no native matches. Interestingly, the single haplotype that was found to dominate the invasive range

was not observed in our native range samples. This is potentially the result of limited sampling from the native range, but could indicate the relative rarity of this haplotype in the native range. In 6/8 of similar studies, the common invasive range haplotype was observed in the native range (Table 1). In either case, the lack of a native range match for this common haplotype precludes our establishing its source population.

4.3 Implications

Knowledge of genotypic diversity, source(s) and genotype distribution is important for controlling invasive plants, especially if there is preliminary evidence that control agents have differential success across multiple genotypes (Charudattan 2005; Gaskin et al. 2005; Morin et al. 2006). Efforts to control Japanese climbing fern are underway, but previous research has focused on other closely related *Lygodium* species. Our finding of low genetic diversity overall and a single haplotype dominating the invasion suggest that a biocontrol solution for *L. japonicum* might be possible. The strategy for finding a biocontrol agent usually involves searching within the native range of the invasive species for natural enemies that are both host specific and damaging to the invasive species (Roderick and Navajas 2003). In the case of the Japanese climbing fern, work on the Old World climbing fern (*Lygodium microphyllum*) serves as an example. Goolsby et al. (2004) identified several genotypes of lygodium gall mite (*Floracarus perrapae*) from Australia and Asia. They concluded that mites had differing success on the invasive fern genotype; the mite genotypes that performed best came from regions where the native fern genotypes were most similar to the invasive genotype (Goolsby et al. 2004). In the case of *L. japonicum*, the fact of a low diversity invasion therefore simplifies the process of identifying biocontrol genotypes. Ideally, knowing the source location(s) of the *L. japonicum* invasion would greatly narrow the search for appropriate natural enemy genotypes, and further

work should expand native range sampling to discover the source region(s) for the common genotype observed in the U.S. invasion. Until biocontrol methods are implemented, it is likely that *L. japonicum* will continue to spread and displace native vegetation throughout the southeastern United States.

REFERENCES

REFERENCES

Alsos IG, Lavergne S, Merkel MKF, Boleda M, Lammers Y, Alberti A, Pouchon C, Denoeud F, Pitelkova I, Pușcaș M, Roquet C, Hurdu B-I, Thuiller W, Zimmermann NE, Hollingsworth PM, Coissac E. 2020. The treasure vault can be opened: Large-scale genome skimming works well using herbarium and silica gel dried material. *Plants* 9(4): 432.

Andreasen K, Manktelow M, Razafimandimbison SG. 2009. Successful DNA amplification of a more than 200-year-old herbarium specimen: Recovering genetic material from the Linnaean era. *Taxon* 58(3): 959–962.

Bakker FT, Lei D, Yu J, Mohammadin S, Wei Z, van de Kerke S, Gravendeel B, Nieuwenhuis M, Staats M, Alquezar-Planas DE, Holmer R. 2016. Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an iterative organelle genome assembly pipeline. *Biological Journal of the Linnean Society* 117(1): 33-43.

Beck JB, Alexander PJ, Allphin L, Al-Shehbaz IA, Rushworth C, Bailey CD, and Windham MD. 2012. Does hybridization drive the transition to asexuality in diploid *Boechera? Evolution* 66: 985–995.

Bellard C, Cassey P, Blackburn TM. 2016. Alien species as a driver of recent extinctions. *Biology Letters* 12(2): 20150623.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Clement M, Snell Q, Walke P, Posada D, Crandall K. 2002. TCS: estimating gene genealogies. *Parallel and Distributed Processing Symposium* 3: 184.

Charudattan R. 2005. Ecological, practical, and political inputs into selection of weed targets: What makes a good biological control target? *Biological Control* 35: 183–196.

Colautti RI, Lau JA. 2015. Contemporary evolution during invasion: Evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology* 24(9): 1999–2017.

Dlugosch KM, Parker IM. 2008. Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology* 17(1): 431–449.

Erkens RHJ, Cross H, Maas JW, Hoenselaar K, Chatrou LW. 2008. Assessment of age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. *Blumea* 53: 407-428.

Gao L, Wang B, Wang ZW, Zhou Y, Su YJ, Wang T. 2013. Plastome sequences of *Lygodium japonicum* and *Marsilea crenata* reveal the genome organization transformation from basal ferns to core Leptosporangiates. *Genome Biology and Evolution* 5(7): 1403-1407.

Gaskin JF, Zhang DY, Bon MC. 2005. Invasion of *Lepidium draba* (Brassicaceae) in the western United States: Distributions and origins of chloroplast DNA haplotypes. *Molecular Ecology* 14(8): 2331–2341.

Genton BJ, Shykoff JA, Giruad T. 2005. High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Molecular Ecology* 14: 4275–4285.

Google LLC. 2021. Google Earth Pro (Version 7.3.4).

Goolsby JA, Makinson JR, Hartley DM, Zonneveld R, Wright, AD. 2004, Pre-release evaluation and host range testing of *Floracarus perrepae* (Eriophyidae) genotypes for biological control of Old World climbing fern. *XI International Symposium on Biological Control of Weeds* 113-116.

Haufler CH, Pryer KM, Schuettpelz E, Sessa EB, Farrar DR, Moran R, Schneller JJ, Watkins JE, Windham MD. 2016. Sex and the single gametophyte: Revising the homosporous vascular plant life cycle in light of contemporary research. *BioScience* 66(11): 928–937.

Holt SD. 2022. What is *Salvinia molesta*? Determining the genetic composition and number of origins of the invasive giant Salvinia. Unpublished.

Hufbauer RA, Sforza R. 2007. Multiple introductions of two invasive Centaurea taxa inferred from cpDNA haplotypes. *Diversity and Distribution* 14: 252–261.

Hutchinson JT, Langeland KA. 2010. Review of two non-native, invasive climbing ferns (*Lygodium japonicum* and *L. microphyllum*), sympatric records and additional distribution records from Florida. *American Fern Journal* 100(1): 57–66.

Kartesz JT. 2015. The Biota of North America Program (BONAP). *Floristic Synthesis of North America*.

Kates HR, Doby JR, Siniscalchi CM, LaFrance R, Soltis DE, Soltis PS, Guralnick RP, Folk RA. 2021. The effects of herbarium specimen characteristics on short-read NGS sequencing success in nearly 8000 specimens: Old, degraded samples have lower DNA yields but consistent sequencing success. *Frontiers in Plant Science* 12: 1076.

Katoh, K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30: 3059–3066.

Kurose D, Pollard KM, Ellison CA. 2020. Chloroplast DNA analysis of the invasive weed, Himalayan balsam (*Impatiens glandulifera*), in the British Isles. *Scientific Reports* 10(1).

Leichty E, Carmichael B, Platt W. 2011. Invasion of a southeastern pine savanna by Japanese climbing fern. *Castanea* 76(2): 293–299.

Leigh JW, Bryant D. 2015. POPART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* 6(9): 1110-1116.

Lott MS, Volin JC, Pemberton RW, Austin DF. 2003. The reproductive biology of the invasive ferns *Lygodium microphyllum* and *L. japonicum* (Schizaeaceae): Implications for invasive potential. *American Journal of Botany* 90(8): 1144–1152.

Madeira PT, Pemberton RW, Center TD. 2008. A molecular phylogeny of the genus *Lygodium* (Schizaeaceae) with special reference to the biological control and host range testing of *Lygodium microphyllum*. *Biological Control* 45(3): 308–318.

McCulloch, GA, Hereward, JP, Lake EC, Smith MC, Purcell MF, Walter GH 2018. The complete chloroplast genome of the invasive fern *Lygodium microphyllum* (Cav.) R. Br. *Mitochondrial DNA Part B* 3(2): 746–747.

McIvor L, Maggs CA, Provan J, Stanhope MJ, 2001. rbcL sequences reveal multiple cryptic introductions of the Japanese red alga *Polysiphonia harveyi*. *Molecular Ecology* 10(4): 911–919.

Minogue PJ, Jones S, Bohn KK, Williams RL. 2009. Biology and control of Japanese climbing fern *(Lygodium japonicum). EDIS* 2009(5).

Morin L, Evans KJ, Sheppard AW. 2006. Selection of pathogen agents in weed biological control: critical issues and peculiarities in relation to arthropod agents. *Australian Journal of Entomology* 45(4): 349–365.

Munger GT, 2005. Lygodium spp. Fire Effects Information System. USA: USDA Forest Service.

Oduor AM, Gómez JM, Herrador MB, Perfectti F, 2015. Invasion of *Brassica nigra* in North America: distributions and origins of chloroplast DNA haplotypes suggest multiple introductions. *Biological Invasions* 17(8): 2447–2459.

Pemberton RW, Ferriter AP. 1998. Old World Climbing Fern (*Lygodium microphyllum*), a dangerous invasive weed in Florida. *American Fern Journal* 88(4): 165–175.

Pimentel D, Lach L, Zuniga R, Morrison D. 2000. Environmental and economic costs of nonindigenous species in the United States. *BioScience* 50(1): 53–65.

RStudio Team. 2021. RStudio: Integrated Development for R. RStudio.

Roderick GK, Navajas M. 2003. Genes in new environments: Genetics and evolution in biological control. *Nature Reviews Genetics* 4(11):889–899.

Saeidi S, McKain MR, Kellogg EA. 2018. Robust DNA isolation and high-throughput sequencing library construction for herbarium specimens. *Journal of Visualized Experiments* 133: e56837.

Saltonstall K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *Proceedings of the National Academy of Sciences* 99(4): 2445-2449.

Steele RP, Pires CJ. 2011. Biodiversity assessment: State-of-the-art techniques in phylogenomics and species identification. *American Journal of Botany* 98(3): 415–425.

Shirk RY, Hamrick JL, Zhang C, Qiang S. 2014. Patterns of genetic diversity reveal multiple introductions and recurrent founder effects during range expansion in invasive populations of *Geranium carolinianum* (Geraniaceae). *Heredity* 112(5): 497–507.

Simon A, Britton R, Gozlan R, van Oosterhout C, Volckaert FAM, Hänfling B. 2011. Invasive cyprinid fish in Europe originate from the single introduction of an admixed source population followed by a complex pattern of spread. *PLoS ONE* 6(6).

Soltis DE, Soltis PS. 1992. The distribution of selfing rates in homosporous ferns. *American Journal of Botany* 79: 97–100.

Staats M, Cuenca A, Richardson JE, Ginkel RV van, Petersen G, Seberg O, Bakker FT. 2011. DNA damage in plant herbarium tissue. *PLoS ONE* 6(12).

Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany* 99(2): 349–364.

Vilà M, Espinar JL, Hejda M, Hulme PE, Jarošík V, Maron JL, Pergl J, Schaffner U, Sun Y, Pyšek P. 2011. Ecological impacts of invasive alien plants: A meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters* 14(7): 702–708.

Williams DA, Overholt WA, Cuda JP, Hughes CR 2005. Chloroplast and microsatellite DNA diversities reveal the introduction history of Brazilian peppertree (*Schinus terebinthifolius*) in Florida. *Molecular Ecology* 14(12): 3643-3656.

Wyatt R, Wyatt GE. 2013. *Lygodium japonicum* (Thunberg) Swartz in the Piedmont of Georgia. *American Fern Journal* 103(3): 188–190.

Zeng CX, Hollingsworth PM, Yang J, He ZS, Zhang ZR, Li DZ, Yang JB. 2018. Genome skimming herbarium specimens for DNA barcoding and phylogenomics. *Plant Methods* 14(1).