Fall 2022

Applying Conservation Genomic Techniques to Guide Management of the Reticulated Flatwoods Salamander (Ambystoma bishopi)

Miranda Gaupp

APPLYING CONSERVATION GENOMIC TECHNIQUES TO GUIDE MANAGEMENT OF THE RETICULATED FLATWOODS SALAMANDER (*AMBYSTOMA BISHOPI)*

by

MIRANDA GAUPP

(Under the Direction of James H. Roberts)

ABSTRACT

The Reticulated flatwoods salamander (*Ambystoma bishopi*) is a federally endangered amphibian endemic to the longleaf-pine ecosystem of the southeastern U.S. This study used analyses of single-nucleotide polymorphism (SNP) data, collected from 2,255 unique individuals across 5 breeding seasons, spread across the known extant range of *A. bishopi*, to characterize the genetic diversity and demographics of populations, genetic relationships among populations, and patterns and spatial extents of gene flow, and to evaluate potential effects of management on *A. bishopi*'s resiliency. Population structure was strongly hierarchical, with individual breeding ponds (n = 38) acting as semi-connected subpopulations within five regional metapopulations (Mayhaw in Georgia; Oglesby, Eastbay, Garcon, and Escribano in Florida). Likewise, gene flow among populations was scale-dependent: negligible genetic differentiation, indicative of high gene flow, was observed only between pairs of ponds separated by < 0.5 km, whereas between 0.5 and 5 km I observed steep genetic isolation by distance, and beyond 5 km genetic differentiation was generally high and only weakly related to distance. Across several breeding seasons, the effective number of breeders (N_b) per pond per year averaged 26 individuals (range 4 to 104). Larger-area, slower-drying ponds located closer to other occupied ponds exhibited larger N_b and greater genetic diversity. Based on genetically-reconstructed pedigrees, the ongoing headstarting program at Escribano successfully captured 97.9% of the estimated total number of alleles, but only 63% of the total number families, in each cohort. Based on these results, I recommend the following: 1) Given its genetic distinctiveness, Georgia populations

merit elevated priority for protection and restoration. 2) Resiliency and redundancy (*a la* the species' recovery plan) should be assessed at the spatial grain of individual breeding ponds. 3) Attempts to restore habitat connectivity should consider dispersal over distances > 500 m to be relatively unlikely. 4) Finally, to the extent that headstarted individuals are used to augment existing or introduce new populations, managers should consider the potential risks of founder effects, and reduce these risks by creating genetically and demographically diverse headstart samples, for example by maximizing the diversity of egg/larva collections over time and space within ponds.

INDEX WORDS: Genomics, Population genetics, Conservation genetics, Endangered species, Amphibians, Ambystoma bishopi, Reticulated flatwoods salamander

APPLYING CONSERVATION GENOMIC TECHNIQUES TO GUIDE MANAGEMENT OF THE RETICULATED FLATWOODS SALAMANDER (*AMBYSTOMA BISHOPI)*

by

MIRANDA GAUPP

B.S., Millsaps College, 2019

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

APPLYING CONSERVATION GENOMIC TECHNIQUES TO GUIDE MANAGEMENT OF THE RETICULATED FLATWOODS SALAMANDER (*AMBYSTOMA BISHOPI)*

by

MIRANDA GAUPP

Major Professor:	James Roberts
Committee:	Scott Harrison
	Lance McBrayer
	Carola Haas (Non voting)

Electronic Version Approved:
December 2022

ACKNOWLEDGMENTS

Thank you to everyone, listed and not, that helped me in pursuing this degree, assisting my growth as a scientist, and bettering this thesis and research. I am extremely proud to present the work that I have accomplished during my time at Georgia Southern University, and I am incredibly grateful to everyone that helped to make it all possible.

To my parents, thank you for your support, encouragement, and understanding during this challenging time in my life. Your words reassured me that I am exactly where I wanted to be during the difficult moments.

To my roommates, Sabina and Noel, thank you for your encouragement, reminders to stop and smell the roses, and giving me a shoulder to cry on. Your friendship means the world to me.

This thesis and research would not have been possible without the help of the many collaborators. Thank you! Especially Charlie Abeles and Kelly Jones and all of the many biologists and techs in the field at Escribano and Eglin for their work in collecting and coordinating samples and sharing their knowledge of this species.

Thank you to my committee members for all of your input, guidance, and support during the course of my graduate studies. To Carola Haas, thank you for sharing your knowledge of this species and for aiding my understanding of this study system.

To my advisor, Jamie Roberts, thank you for your unwavering patience, support, and understanding during these three years. Your mentorship has allowed me to grow immensely as a scientist, and I am incredibly grateful for all you have done for me. If it were possible, I would gladly continue to work under your guidance as a Ph.D. student.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
GENERAL INTRODUCTION

Study Overview

The reticulated flatwoods salamander (*Ambystoma bishopi*) is endemic to mesic flatwoods in the longleaf-pine ecosystems of the southeastern U.S. (Palis, 1997a; Bevelhimer et al. 2008). The historic range of the species includes areas in south Alabama, north Florida and South Georgia west of the Apalachicola-Flint rivers (IUCN 2020). The majority of published research on *A. bishopi* has focused on populations occupying Eglin Air Force Base with limited range-wide studies of the species outside of phylogenetics (but see Williams et al., 2021). These studies are informative but provide limited information on range-wide diversity and genetic relationships of the species. Collecting and analyzing data from other known populations of *A. bishopi* can provide a more thorough understanding of the species' status and ecology through estimates of effective population size, relatedness indices, genetic diversity statistics, and gene flow patterns. Recently developed genomic sampling techniques allow examination of thousands of markers from many thousands of individuals for a low cost relative to genetic approaches. This economical approach allows us to deeply sample individuals and populations as well as widely sample the range of the species in order to assess genomic diversity and relationships of multiple populations. This study uses genomic samples from Mayhaw Wildlife Management Area (Mayhaw) and Garcon Point Water Management Area (Garcon) as well as extensive sampling from both Eglin Air Force Base (Eglin) and Escribano Point Wildlife Management Area (Escribano) properties to assess range-wide genomic diversity and relationships. By applying genomic techniques to samples encompassing a broader range of the species, knowledge gaps in *A. bishopi* evolutionary dynamics can be filled, such as delineating gene flow patterns and estimating genetic differentiation between and among populations, as well as

estimating effective population sizes and genetic diversity levels for populations across the species range. This information can then be used to guide management decisions to better conserve the species. The permeability/resistance of habitat connecting populations influences dispersal rates and gene flow within the metapopulation, consequentially impacting the allelic richness of each population and the genetic differentiation between populations. Landscape genetics information such as this can further guide management decisions to maintain population connectivity and gene flow in order to preserve genetic resiliency. Additionally, estimates of genetic diversity and differentiation can direct management efforts to preserve highly diverse populations in order to maintain allelic diversity within the metapopulation as a whole (Funk et al. 2012; de Guia and Saitoh 2007).

A key first step to investigating population dynamics is an accurate understanding of population structure, which can be delineated at very fine scales using population genomic techniques. This can be done by using genomic estimates of relatedness to visualize how individuals are clustering into populations. By visualizing clustering patterns using STRUCTURE (Pritchard et al. 2000; Pritchard, Wen, and Falush 2010) and principal coordinates analysis (PCoA), we can see how the population is spatially structured across the landscape and which local breeding groups (i.e. ponds) are more or less related to each other. This information sheds light on gene flow patterns between ponds. Ponds that cluster more closely together have lower genetic differentiation and therefore more gene flow between them as opposed to ponds clustering farther apart. Environmental conditions such as distance or landcover can be used to characterize these clustering patterns to identify habitat conditions that influence functional connectivity.

Once populations are delineated, population demographics and genetic diversity can be characterized, including the effective number of breeders, observed and expected heterozygosity, allelic richness and various indices of relatedness and family structure. These population demographic and genetic estimates allow more accurate assessments of population resiliency and guide management decisions by indicating populations in need of potential genetic rescue. This allows conservation activities to focus on restoring populations of concern and preserving healthier populations. Demographic and genetic estimates can also be related to environmental data to identify environmental conditions that promote larger, more diverse populations. A microsatellite study on *A. bishopi* metapopulation ecology on Eglin found a negative correlation between genetic diversity and distance to other occupied ponds, and a positive correlation between pond area of suitable breeding habitat and genetic diversity (Wendt et al., 2021). This study also found low and variable effective population sizes of ponds at Eglin. These results highlight the importance of maintaining population connectivity and suitable habitat to preserve genetic diversity, and thus population resiliency. By characterizing genetic diversity and demographics for populations across a larger extent of the species range, population resiliency can be more accurately assessed for a greater number of populations and for the species as a whole.

Headstarting programs have been ongoing at Eglin and Escribano to maintain population census numbers, but the genetic impacts of headstarting cohorts is unknown. If headstarting cohorts capture a small fraction of genetic diversity relative to the pond population of origin, captively rearing highly related individuals and releasing them back into source ponds could decrease the population's overall genetic diversity, cause a genetic bottleneck, increase inbreeding risk, and decrease adaptive potential (Ryman and Laikre, 1991; Laikre et al., 2010;

Weeks et al., 2011). However, if captively reared individuals adequately capture the genetic diversity of their origin pond, headstarting could successfully supplement the population's demography while maintaining genetic diversity. By targeting a panel of neutral SNPs in a time series of headstarted individuals alongside naturally reared individuals, we can evaluate how well headstarting efforts are sustaining population genetic diversity through estimates of the effective number of breeders, allelic richness, heterozygosity, and relatedness.

A "final frontier" of my study is to identify genome regions under natural selection. This adaptive genomic information allows us to 1) determine compatible donor-recipient ponds for translocation and reintroduction events based on local adaptation and 2) strategically ensure the maintenance of important alleles when designing headstart programs. We can do this through a genotype-environment approach (GEA) which allows us to determine genotypes based on environmentally adapted genome regions and match ponds which select for the same genotypes. Determining compatible ponds through identifying adaptive genomic regions allows us to avoid outbreeding depression and the swamping out of local adaptations when translocating headstarted individuals. By using both neutral and adaptive markers to identify source-recipient pairs for translocation and reintroduction, we have a lower likelihood of moving an individual to an unsuitable pond and thus a higher likelihood of increasing overall population diversity.

With my thesis, I use population and landscape genomic techniques to enhance our range-wide understanding of genomic relationships within and between populations of *A. bishopi* and evaluate in detail the metapopulation dynamics, genetic diversity, and conservation status of the Escribano and Eglin metapopulations of this species. My overall goals are to 1) better understand the genetic structure and relationships of *A. bishopi* populations, 2) assess resiliency through characterizing populations demographics and genetic diversity, 3) evaluate the efficacy

of headstarting programs, 4) evaluate how habitat conditions regulate connectivity, and 5) ultimately, provide guidance on how to more effectively conserve this species and its habitat.

Brief Introduction to Amphibians

Class Amphibia (amphibians) consists of ectotherms characterized by thin, permeable skin which allows for cutaneous respiration in exchange for moisture dependence to avoid desiccation. To avoid desiccation, amphibians regulate body temperature and moisture mostly through behaviors including burrowing and nocturnal activity. As a result of their moisture dependence and ectothermic nature, amphibians are most diverse in warm, humid areas of the world (Buckley and Jetz 2007) though they have adapted to live in a variety of habitats including desert and mountain environments (Mayhew, 1965; Pilliod et al. 2002).

Amphibians have diverse and often complex life cycles (Wilbur 1980; Werner and Gilliam 1984) with one or more aquatic life stages. An amphibian life cycle typically consists of aquatic egg and larval stages, either a terrestrial or aquatic juvenile stage, and an adult stage that is either terrestrial, semi-aquatic, or fully aquatic. The complex and variable life history of amphibians is exemplified in by the life cycle of the Eastern Newt (*Notophthalmus viridescens*) in Figure 1.1. Metamorphosis allows for the transition from aquatic to terrestrial living. After metamorphosis, juveniles may remain aquatic or leave breeding grounds (depending on the species) and continue to develop into adults in the surrounding terrestrial habitat. As adults, amphibians migrate to breeding grounds to reproduce. Breeding grounds can be streams, ephemeral pools, or permanent ponds, and the timing of adult migration to breeding grounds varies between and within species (Grant et al., 2009; Wells, 2010; Arnfield et al., 2012; Brooks et al., 2019). Because amphibians are dependent on both aquatic and terrestrial habitat, as well as

the connecting landscape, they often are especially sensitive to habitat alterations, degradation, and fragmentation (Alford 2010; Guerry and Hunter 2002; Zamudio and Wieczorek 2007).

Amphibians play an integral role in energy transfer within ecosystems across the globe. The critical role of amphibians in food webs can be explained by a) the biphasic life cycle of most amphibians, b) their mid-trophic level class, serving as both predator and prey, and c) the large portion of vertebrate biomass accounted for by amphibians, particularly in forest and wetland ecosystems (Burton and Likens 1975; Gibbons et al., 2006; Hamer and McDonnell 2008; Clipp and Anderson 2014). Because amphibians occupy aquatic and terrestrial habitats, amphibians are critical in connecting energy transfer between these habitats through the exchange of energy and biomass that occurs in migration and dispersal. This energy and biomass exchange provides food sources to the surrounding habitat, therefore a decline in amphibian populations would largely disrupt trophic systems (Wilson & Dorcas, 2003; Ranvestel et al., 2004; Meyer et al., 2007; Halliday, 2008; Peterman, Crawford, & Semlitsch, 2008)

Amphibians facing decline

Amphibians as a group are facing global population declines and imperilment (Fisher and Shaffer 1996; Houlahan et al. 2000; Collins and Storfer 2003; Stuart et al. 2004). About 2,200 of the earth's amphibian species (approximately 31.9%) have been listed as vulnerable or higher on the IUCN red list, with main threats consisting of habitat loss, disease, hazardous chemicals, invasive species, and climate change (Daszak, Cunningham, & Hyatt, 2003; Davidson 2004; Daszak et al. 2005; Cushman 2006; Collins, 2010; Salice 2012). Because of their permeable skin, amphibians are sensitive to environmental pollutants and habitat alterations, which may contribute to the greater number of population declines and extinctions observed is amphibian species compared to birds and mammals (Stuart et al. 2004; Blaustein et al. 2011; IUCN red list

2022). Conserving habitat in areas with high amphibian diversity, such as the southeastern United States, is necessary in order to preserve species diversity and prevent further decline.

The southeastern United States houses approximately 140 of the country's 295 amphibian species (Bartlett and Bartlett 2006; Graham et al. 2010). Of the 140 species of amphibians in the southeastern U.S., 102 (approximately 72.9%) are salamanders. These 102 species make up 17% of the world's salamander species (Mitchell and Gibbons 2010). Although the southeast United States is a hotspot for salamander diversity, many of these species are narrowly endemic which increases their vulnerability to potential threats such as habitat loss and fragmentation. To conserve these vulnerable amphibians, management decisions must be informed through an understanding of 1) the basic biology and ecology of the group, 2) information on species distribution and range to distinguish critical habitat, and 3) an understanding of gene flow to preserve genetic diversity. However, our understanding is still limited for many salamander species and few species have adequate information to assess conservation status (Barata, Uhlig, Silva,& Ferreira, 2016).

Amphibian metapopulation dynamics

Amphibian breeding wetlands often are considered classic examples of metapopulation structure because they are discrete patches connected through periodic dispersal (Smith and Green, 2005; Marsh and Trenham, 2008; Driscoll 1997; Hels and Nachman 2002). The congregation of individuals into ponds creates populations which could be connected through movement and gene flow to form 1) a panmictic or patchy metapopulation (all ponds are connected through high rates of dispersal), 2) a Levins-style metapopulation (each pond has equal chance of undergoing local extinction/colonization), 3) a mainland-island or source-sink metapopulation (migrants from large, more stable populations disperse to create smaller, less

stable populations), 4) an isolated or non-equilibrium metapopulation (there is no gene flow between ponds/populations), or 5) an intermediate metapopulation which has a mixture of the characteristics of the other metapopulation models (Harrison 1991; Harrison and Hastings 1996). These metapopulation examples can be seen in Figure 1.2.

Because ponds are discrete patches, amphibian metapopulations typically are expected to have 1) population dynamics determined primarily by processes at breeding ponds (ponds-as-patches view; Marsh & Trenham, 2008), 2) extinction/recolonization of subpopulations as common occurrences with local extinctions as a result of stochastic processes, and 3) pond isolation affecting colonization/extinction and occupancy due to limited dispersal ability. These assumptions are important to understand because of their implications for monitoring and managing amphibian populations, however they are not true for every species. Many factors can contribute to the dynamics of an amphibian metapopulation. The condition of the terrestrial habitat surrounding breeding ponds can affect population dynamics in terms of growth and population size in effect of available shelters or habitat quality (Alford, 1996; Loredo et al., 1996; Skelly et al., 1999). Populations in more disturbed habitats may have more limited dispersal due to barriers such as roads (Fahrig et al. 1995; Gibbs 1998) whereas populations in relatively undisturbed habitats may show no significant effects of isolation (Skelly et al., 1999). These aspects influence observed extinction/recolonization patterns (Wells 1977; Semlitsch et al., 1996). Species specific studies must take place to assess the accuracy of these assumptions and adjust management to their implications. For example, movement corridors can be created to combat barriers and facilitate dispersal in more disturbed habitats. In relatively undisturbed habitat, management decisions can be directed to altering landscape in order to optimize habitat conditions and facilitate the recolonization of unoccupied areas.

Ambystoma biology and ecology

Thirty-two of the current 741 described salamander species belong to the family Ambystomatidae. This family is composed of a singular genus, *Ambystoma*. Many of these ambystomatids are pond breeders, seasonally migrating to breeding ponds from the surrounding upland habitat in which they spend most of their adult life underground. This fossorial adult life-strategy presents challenges for field collection outside of their breeding season. As a result, most of what we know of ambystomatids comes from juvenile individuals collected during breeding seasons when adults emerge from underground retreats and congregate in ponds to reproduce (Heyer, Donnelly, Foster, & Mcdiarmid, 1994). Because of this juvenile-based study approach, field estimates of census size may be inflated since many larvae and juveniles will not survive to adulthood thus biasing our understanding of the species' demographic status.

Ambystomatids are widely believed to exhibit limited dispersal over terrestrial habitats due to their small bodies and moisture dependence. Smith and Green (2005) reviewed the literature at the time and reported that 94% of published maximal dispersal distances for salamander species (across 37 species) were < 1 km meaning most individuals may not disperse very far. *Ambystoma maculatum* disperse 300-500 meters (Madison 1997) but mostly stay withing 90 m of their breeding pond (Semlitsch 1998), which was supported through genetic data showing reduction in gene flow between ponds separated by > 4.8 km (Zamudio and Wieczorek 2007). Movement events ≥ 500 m are rare in *Ambystoma bishopi* as found through genomic microsatellite data from individual on Eglin(Wendt et al., 2021). Because of their dispersal limitations, the geographical distance between ponds could have high influence on gene flow patterns as well as recolonization of populations after local extirpation. Besides distance, environmental conditions between ponds, such as soil moisture or prominent vegetation type, may have a large influence on gene flow patterns due to individuals' dependence on moisture.

Assuring that individuals metamorphose and reach the juvenile stage is a critical component of gene flow in *Ambystoma* populations. In most species, juveniles are the dispersing life stage (Gamble et al., 2007) meaning that population diversity and persistence is reliant on this life stage (Taylor et al 2006, Harper et al 2008). Two environmental variables play critical roles in successful metamorphosis: 1) hydroperiod- the length of time water is in a wetland, and 2) recession rate- the rate at which water is lost from a wetland (Semlitsch and Wilbur 1988; Semlitsch 2002; Baldwin et al. 2006; Gomez-Mestre et al. 2013). These two variables determine the period of time in which metamorphosis must occur and the period of time in which larvae can grow and have access to suitable resources such as food and preferred habitat (Chandler et al. 2017). Due to dispersal limitations of *Ambystoma* salamanders and patch-dependency that is sensitive to changes in vegetation and hydrologic conditions, they are an ideal taxon for investigating genetic connectivity within metapopulations (Joly et al., 2001; Marsh & Trenham 2008; Semlitsch and Anderson, 2016).

The Reticulated Flatwoods Salamander

Basic biology and life history

As a pond breeder specializing on discrete, declining habitat types, *A. bishopi* makes a useful case study for understanding both the general metapopulation dynamics and landscape ecology of *Ambystoma* and the ways in which habitat alterations and management programs affect the viability of this and other imperiled species. *Ambystoma bishopi* inhabits the longleaf pine flatwoods ecosystem (Palis, 1997b; Bevelhimer et al. 2008) of the southeastern USA coastal plain (Pauly et al. 2007; IUCN 2020). Due to environmental changes such as fire suppression, land conversion, drought, and loss of habitat, *A. bishopi* populations have been reduced and fragmented such that the combined range of *A. bishopi* and the closely related *A. cingulatum*

decreased from 476 locations recorded prior to 1999 to 63 locations recorded between 2010 and 2015 (Bevelhimer et al., 2008; Bishop and Haas 2005; Chandler et al. 2016; Pauly et al., 2012). This population reduction is depicted by the Semlitsch et. al (2017) range map (Figure 1.3). Other factors that pose threats to *A. bishopi* status are invasive plant species, climate change, small population sizes, pesticides and herbicides, feral pigs, and predatory fish presence in breeding ponds (Walls et al. 2013; Chandler 2015; Duellman and Trueb, 1986; Semlitsch 1987). *A. bishopi* is listed as federally endangered by U.S. Fish and Wildlife Service (USFWS, 2009) and vulnerable on the IUCN Red List (Palis & Hammerson, 2008; IUCN 2020). Twenty population centers were known as of 2009; six of these occurred on public lands and were known to be occupied in 2014, but most of these sites exist on private land and cannot be regularly surveyed (U.S. Fish and Wildlife Service species review and evaluation 2014; Farmer et al. 2016; O'Donnell et al. 2017). Of these 20 population centers, 2 were located in southwest Georgia and 18 were in Florida. Fourteen population centers consisted of a single known breeding pond with no confirmed occurrences within the past two decades (Pauly et al. 2007; Semlitsch et al. 2017). Two of the largest, most consistently occupied strongholds for *A. bishopi* are on Eglin Air Force Base (Eglin) and Escribano Point Wildlife Management Area (Escribano). Two other known population centers that are significantly smaller are located on Mayhaw Wildlife Management Area (Mayhaw) and Garcon Point Water Management Area (Garcon).

A. bishopi is a moderately sized species for its genus, with adult length averaging 5.5 inches and females having larger adult size over males. Sexual maturity typically is reached after 1 year of age for males and 2 years for females (Palis, 1997b). The lifespan of *A. bishopi* is still uncertain due to difficulties in tracking individuals in the field through the duration of their life, but they are estimated to live up to 13 years in the wild (Brooks et al 2020). Consistent with

other amphibians, *A. bishopi* has a complex life cycle including terrestrial adult, juvenile, and egg stages and an aquatic larval stage. Unlike most ambystomatids, *A. bishopi* lays eggs terrestrially in dry beds of ephemeral breeding ponds (Anderson and Williamson, 1976). During late fall and early winter months (October through December), adults emerge from upland habitat and migrate to dry breeding ponds to court and deposit eggs in micro-depressions among herbaceous vegitations. Eggs begin to develop immediately but only hatch when inundated during winter rains, usually occurring between December and February (Anderson and Williamson, 1976; Palis, 1995; Palis, 1997b; Bevelhimer et al., 2008). This strategy may elevate the risk of clutch failure due to poorly timed rainfall (Martin, 1999).

Larval development is followed by metamorphosis into juveniles which is believed to be heavily influenced by environmental factors. For example, metamorphosis is thought be initiated by pond recession (Palis, 1995). Duration of hydroperiod, depth of pond, and recession rate are considered important factors for successful larval development and metamorphosis due to their influence on the availability and abundance of *A. bishopi* larval prey items, accessibility to preferred habitat within the pond, and length of time larvae have to grow and develop (Chandler et al., 2015; Chandler et al., 2017). The larval size at which metamorphosis occurs varies between individuals and is dependent on hydroperiod (Brooks et al., 2020). Longer hydroperiods result in larvae postponing metamorphosis in favor for a continued growth period (Amburgey et al 2012; Brooks et al 2020). Longer hydroperiods, increased larval growth periods, and increased size at metamorphosis may increase lifetime survival rate, which the underscores importance of these factors in considering the ecology and evolution of the species (Wilbur & Collins 1973; Semlitsch, Scott, & Pechmann, 1988; Scott 1994; Kingsolver & Pfenning 2007; Cabrera-Guzmán et al., 2013). These pond characteristics may influence local adaptation of *A. bishopi*

populations (Richardson & Urban, 2013; Giery et al., 2021), but further study is needed on the topic.

After metamorphosis, juveniles emerge from breeding ponds and disperse throughout the surrounding mesic upland habitat. These juveniles continue to grow during their adult life stage but tracking growth during this stage is difficult due to the fossorial nature of adults (Brooks et al 2020). Adults seek moist refuges such as underground burrows, crayfish holes, and decomposing logs to avoid desiccation. Above-ground movement is rare outside of breeding migrations. However, environmental and endogenous cues of temperature, precipitation, and the day of the year are strongly tied to above-ground movement of *A. bishopi*, supporting preference for moderate temperatures and moisture/precipitation. Few adults move around above ground before October 25th or after February 25th (Brooks et al 2019), but dispersal of metamorphs typically occurs between late March and May (Chandler et al. 2017; Haas, pers. comm.).

Metapopulation dynamics of Ambystoma bishopi

The moisture-dependency and small-body dispersal limitations of *A. bishopi* suggest important roles of both environmental conditions and between-pond distance in shaping metapopulation structure. The influence of environmental conditions and distance was supported in a study by Brooks et al. (2019) through occupancy-based metapopulation models and eigenvector mapping. The results indicated a decrease in pond connectivity with increased spatial isolation. Ponds separated by distances of 1.5 km or more were observed to have no demographic interaction. Using population genetic methods to infer dispersal, Wendt et al. (2021) found movements >500 m to be highly infrequent. Brooks et al. (2019) found that the amount of herbaceous vegetation in ponds best predicted *A. bishopi* occupancy status as compared to wetland area, maximum wetland depth, proportional area of breeding habitat, and

average hydroperiod, which suggests that environmental factors influence the occupancy status of an area. The ephemeral nature of *A. bishopi* breeding ponds and possible elevated clutch failure caused by terrestrial egg laying provides support for a Levins style metapopulation, where each population has a nontrivial chance of local extinction (the disappearance of a population (breeding pond)) each season (Harrison, 1991). In order for extinction-recolonization patterns and demographic rescue effects between breeding ponds to persist and maintain the metapopulation dynamic, pond connectivity must be properly understood.

A. bishopi's fossorial adult-life strategy presents difficulty for understanding its population structure and dynamics within and between local breeding populations. However, recent studies on this species have begun to fill key information gaps, including 1) hydroperiod influence on metamorphosis, 2) population size estimates, 3) landscape connectivity, 4) growth rate, 5) factors related to pond occupancy, and 6) environmental drivers of migration (Anderson and Williamson ,1976; Wendt et al., 2021; Brooks et al., 2019; Chandler et al., 2016; Brooks et al., 2020). The majority of ecological studies and published literature on *A. bishopi* focuses on populations on Eglin. Collecting environmental, demographic, and genetic data from other populations of *A. bishopi* and applying population and landscape analyses can provide a more thorough understanding of the genetic structuring of populations, effective populations sizes, and dispersal patterns among others to better our knowledge of the species status and ecology. Filling gaps in our understanding of *A. bishopi* ecology and continuing to monitor its status is vital in our effort to guide management and aid in recovery of the species.

Management activities for Ambystoma bishopi

Current management activities including habitat management, population monitoring, and headstarting are in place at Escribano and Eglin to conserve this species and potentially

employ headstart individuals in translocation and reintroduction efforts in order to 1) augment occupied breeding populations and 2) repopulate suitable, yet unoccupied breeding habitat. Habitat management includes manual and chemical removal of trees, alongside regular controlled burning to clear large woody vegetation and accumulated litter and preserve the longleaf pine ecosystem through germinating herbaceous ground cover. Removal of trees is essential to restoring habitat for *A. bishopi* because pine tree overgrowth a) reduces food production through shading ponds and b) shortens pond hydroperiods through increasing evapotranspiration. Population monitoring takes place in the form of annual surveys for presence of *A. bishopi* through mostly flashlight searches and dipnet surveys.

Employing genetic monitoring approaches such as estimating effective breeding population size (N_b) and genetic diversity statistics would allow a more thorough and accurate understanding of population demographic status and genetic viability and provide valuable information on species and population status that is unattainable through egg/larval capture data. For example, estimating N_b for breeding ponds could provide a better understanding of 1) species demographics through estimates of how many adults are involved in producing the generation under study, 2) genetic viability through insight on the rate at which the population of study experiences loss of genetic material, and 3) population structure through the rate at which the population experience genetic drift (Charlesworth et al., 2009; McCartney and Shaffer, 2018; Murphy et al., 2018; Shaffer et al., 2015; Martinez-Solano & Gonzalez, 2008). Genetic monitoring practices also provide a database to track species and population genetic and demographic trends over time. This genetic information could be coupled with landscape and environmental data to relate environmental occurrences, such as periods of drought, to changes in diversity. Additionally, genetic monitoring techniques could provide information on the

genetic and demographic impact of headstarting programs since this area is poorly understood. In all these ways, applying genetic techniques to monitoring *A. bishopi* populations could advance our understanding of population demographic and genetic status.

Headstarting programs have been a management method at Eglin and Escribano in order to preserve *A. bishopi* populations. It is widely believed that amphibians are most vulnerable to threats such as disease and predation during pre-adult life stages of egg, larvae, and metamorph. This belief partly motivated the rise of amphibian headstarting programs in which early amphibian stages of eggs and/or larvae are collected and reared in aquaria (cattle watering tanks) until the late metamorph or juvenile stage (IUCN-SSC, 2013). After reaching an appropriate developmental stage, headstarted cohorts are released back into the breeding ponds from which they were collected in order that they disperse naturally into the surrounding habitat. Headstarting programs usually are initiated with hopes of increasing early-life stage survival, and thus population size and stability (Quinn, 1980; Skriver, 1988; Banks, Beebee, and Denton, 1993). The genetic diversity captured in headstarted cohorts are rarely surveyed to understand the heterogeneity being released into the population after rearing. It is a popular management and conservation strategy, but the demographic and genetic effects of captive rearing should be explored in order to understand it's impact and avoid negative consequences such as decreased effective population size or increased inbreeding levels (Wang and Ryman, 2001; Araki et al., 2007).

Translocation and reintroduction of captive-reared individuals is sometimes pursued in hopes of increasing genetic diversity of populations or repopulating suitable but unoccupied habitat, and there is potential for future use of this tactic with *A. bishopi*. Two approaches to determine suitable translocation/reintroduction operations using genomic information are 1)

attempt to mimic “natural”, historical dispersal/gene flow patterns, and 2) avoid the mixing of separate gene pools that possess different local adaptations (Moritz, 1999). Without an understanding of potential local adaptation to pond-specific environmental factors and the amount of genetic diversity captured in captive cohorts, translocation and reintroduction of headstarted individuals could cause negative effects like outbreeding depression and genetic swamping (Frankham et al., 2011; Huff et al., 2011). Because of the influence of recession rate and hydroperiod on reproductive and developmental success, selection may be acting according to these variables and should be explored before attempting translocation or reintroduction. Before translocation or reintroduction is attempted, we also must gain an understanding of current genetic status for each pond. Estimates of heterozygosity levels, relatedness, and N_b could be used to identify ponds in most need of genetic rescue, and ponds that are suitable source populations (Zeisset and Beebee, 2013).

Population, Conservation, and Landscape Genomics

Population genomics

The preceding paragraphs outlined key information gaps for recovery of *A. bishopi*, many of which can hopefully be resolved through the application of modern molecular genetic tools. In particular, I have applied population-, conservation-, and landscape-genetic analyses to rich, genome-wide molecular marker data to address the goals of my thesis. Population genomic studies use thousands of gene markers throughout the genome to gain an understanding of population genetic diversity and differentiation and obtain a holistic picture of demographic and evolutionary processes affecting the genome at the population level. Sequentially, the number and location of data markers used for data analysis distinguishes population genomics from population genetics. Population genetic studies generally consist of <30 selectively neutral loci,

whereas population genomic studies generally use thousands of markers from throughout the genome and can target neutral and/or adaptive genome regions depending on the research question of interest (Luikart et al. 2018).

Genomic assessments of population status and interactions has led to a more accurate understanding of population resiliency and evolutionary processes to inform conservation efforts surrounding these topics (Peirson et al., 2016; Allendorf, Hohenlohe, & Luikart 2010; Bernatchez et al., 2017; Shafer et al., 2015). For instance, genomic techniques are thought to provide more reliable measures of inbreeding and inbreeding depression (Kardos, Luikart, & Allendorf, 2015; Kardos et al., 2016; Wang, 2016) and more accurate estimates of population size which in turn provide information on population health and status. For example, results from Camacho-Sanchez et al. (2020) comparison of genetic diversity estimates using microsatellite and SNPs suggest the larger number of loci in SNP studies provide more reliable estimates heterozygosity and population structure. Using this reasoning, a study applying SNPs can more reliably estimate genetic diversity and population structure for *A. bishopi* populations compared to past microsatellite estimates of Wendt et al. (2021). This would in turn provide a more accurate depiction of pond connectivity for example, which could then be used to inform translocation and reintroduction efforts by identifying pond source and recipient pairs.

Amphibians present interesting and challenging cases for population genomic studies due to their large genomes. Amphibians have one of the largest genomes among vertebrates (Sessions, 2008) with a range from ~15 gigabases (Gb) to ~120 Gb, whereas most mammals have 3-4 Gb genomes (Kapusta et al., 2017; Gregory 2018). The size of these genomes is mostly attributed to the large number of long terminal repeat retrotransposons (Sun et al 2012; Sun and Mueller 2014) and long introns which possibly have a greater number of regulatory regions than

mammalian genomes (Smith et al 2009; Nowoshilow et al 2018). The size of the *Ambystoma* genome is estimated to be ~30 Gb based on the *Ambystoma mexicanum* genome which is 32 Gb (Nowoshilow et al. 2018). With this in mind, we can presume *A. bishopi* has a comparably sized genome which presents challenges in successful sequencing and analysis including insufficient depth of coverage, limited number of individuals that can be sequenced per sequencing lane, and slowed computational processing and analysis (Guo et al. 2012; Weisrock et al., 2018).

The challenges presented by the large genome of *A. bishopi* may be best matched through combining sequencing methods, namely restriction site associated DNA sequencing (RADseq) and sequence capture methods. The combination of these sequencing methods is termed RADcap (restriction-site-associated DNA capture) (Hoffberg et al., 2016). RADseq subsets the genome for sequencing through restriction enzyme fragmentation followed by size selection to produce thousands of anonymous loci from throughout the genome (Miller et al 2007). In RADcap ,once RAD laboratory protocol is complete, sequence capture methods begin. Sequence capture uses oligonucleotide probes to selectively target specific RAD loci regions for subsequent amplification and sequencing (Gnirke et al 2009). Through combining RADseq with sequence capture in RADcap methods, the number of fragments produced by RADseq is decreased to only the genomic areas of interest thereby reducing the sequencing "noise" caused by noninformative genome regions. Through applying RADcap methods, target loci can be reliably sequenced across many more individuals at once than with RADseq or sequence capture alone, while still producing deep sequencing coverage of each individual (Hoffberg et al., 2016; Ali et al., 2016; Komoroske et al., 2019). Taking all of this into consideration, RADcap presents itself as a method which can meet the challenge of a large genome species and allow for a more reliable

and in-depth population genomic analysis through high throughput sequencing of thousands of *A. bishopi* individuals.

A typical genome-scale study will uncover a combination of a small number of loci that appear to be under selection and a large number of loci that appear to be selectively neutral; each type is useful for answering different sets of questions. Neutral markers are useful for questions related to genetic drift and gene flow including migration patterns, dispersal, effective population size, relatedness indices, neutral population genetic structure, neutral genetic diversity, and neutral genetic differentiation (Slate et al., 2004, Luikart et al., 2018). Outcomes of these studies can be used to evaluate extinction risk and ability to respond to environmental change (Frankham, 2005; Reed and Frankham, 2003) which can further inform on overall population genetic health and resiliency. Conservation units (within-species units used to help guide management and conservation decisions) can be delineated using neutral genome markers (Funk et al. 2012) as a step to increase population growth and to monitor population status. Neutral genomic markers can be applied in *A. bishopi* populations to estimate genetic and demographic characteristics and assess the impacts of headstarting programs.

Adaptive markers are useful for answering questions related local adaptation and genetic differentiation caused by adaptation. Genomic regions affected by local adaptation are being selected due to local environmental conditions. These adaptive regions can drive gene flow and genetic differentiation within a metapopulation which would not be detected when analyzing neutral loci alone (Hoffmann et al., 2003; Chapman et al., 2009). Adaptive genomic data makes it possible to investigate environmental influences on the genome in species of concern without risking invasive manipulative experiments (Hoffmann & Sgro, 2011; Harrisson et al., 2014; Hoffmann et al., 2015). Insight on local adaptation can inform conservation actions in order to

improve population's adaptive capacity and evolutionary potential through informing and improving 1) restoration by maintaining wetland environmental conditions that shape adaptive loci, thereby conserving genetic diversity (Petranka & Holbrook 2006), 2) reintroduction efforts by matching genotypes to environmental conditions, and 3) translocation events by matching source-recipient pond pairings based on adaptively significant environmental conditions in order to avoid outbreeding. Failure to account for local adaptation can decrease success of reintroduction and translocation because individuals lack the selected traits to thrive in the environmental conditions (Sagvik, Uller, and Olsson 2005), and cause a misunderstanding of gene flow and metapopulation structure (McKay & Latta 2002; Semlitsch 2008). Management of metapopulations without an understanding of local adaptation can lead to a loss of genetic diversity and the associated risks such as population decline and local extinction (Stockwell et al., 2003). Investigation of locally adaptive genome regions in *A. bishopi* populations can inform reintroduction and translocation plans of captively reared individuals to maintain species diversity as a whole. Understanding local adaptation can also benefit *A. bishopi* through identifying breeding ponds of adaptive significance and prioritizing these populations in conservation planning.

Conservation Genomics

Conservation genomics is a subfield of population genomics in which the application of population genomic methods and theory is used to answer questions about the genetic composition and viability of imperiled populations and species. Genome sequencing techniques have shed light on the biology of wildlife species to aid in conservation management through demographic analyses, estimation of genetic variation associated with local adaptation and

fitness, evaluation of inbreeding levels, and detection of loci associated with disease (Larsson et al., 2008; Hoglund et al., 2011; McMahon et al., 2012; Strand et al., 2012).

The effective population size (N_e) is the size of an ideal population that has the same rate of change of heterozygosity as the observed population (Fisher, 1930; Wright, 1931), and is commonly estimated in conservation genomics studies. N_e is used to provide insight into population heterozygosity levels, extinction risk, and viability (Charlesworth, 2009). However, it is difficult to estimate for age-structured species, in which case the effective number of breeders (N_b) may be more suitable. N_b is a similar metric to N_e in that both are used for monitoring population size and genetic diversity. Properly estimating N_e can require waiting several years between sampling events (Waples and Yokota, 2007), whereas N_b can be estimated using samples from a single cohort, which allows for annual monitoring of population status. Annual N_b estimates permit early detections of population declines to help prevent the loss of genetic diversity and population extirpation (Schindler et al., 2010; Schwartz, Luikart, & Waples, 2007). In this way, N_b can be used to monitor population resiliency. In a species such as *A. bishopi* where population decline and loss of genetic diversity is of particular concern, monitoring N_b across many generations can provide valuable information on population resiliency through monitoring trends in population size and genetic diversity.

Conservation biology often seeks to determine population boundaries in order to designate conservation units. Genomic techniques allow estimations of the spatial scale of gene flow and the delineation of population structure both within and between major populations. Delineating gene flow between populations is essential in ensuring a species' sustainability since gene flow contributes to the maintenance of genetic diversity by retaining effective population size, avoiding inbreeding, and sustaining allelic richness and adaptive potential (Frankham 2005;

Reed and Frankham 2003). Using genetic approaches, Blouin, Phillipsen, and Monsen (2010) were able to identify strong differentiation between six major groups of Oregon spotted frog which supports each of these groups as evolutionarily significant in conservation of the species' diversity. These results provide support for management actions to prioritize the conservation and preservation of each of these populations in order to maintain the species overall genetic diversity and resilience. In these ways, genomic data has aided our ability to understand gene flow patterns by providing more power for detecting gene flow and in result delineating populations than previous methods (Gompert et al. 2003; vonHoldt et al. 2011). By applying genomic methods to delineate population structure and gene flow in aggregations of *A. bishopi*, distinct groups of evolutionary and conservation significance can be identified. More specifically, genomic methods can direct management efforts of *A. bishopi* through providing a data-driven method of delineating populations and metapopulations when measuring resiliency, redundancy, and representation.

Genomics has provided the ability to detect loci influenced by selective forces and allow conservation actions to address species' adaptation. This ability enables the identification of genetic changes associated with local adaptation and the environmental conditions influencing fitness. Understanding local adaptation may aid in evaluating populations' potential to respond to environmental changes (Hoffman and Sgro 2011), defining conservation units (Manel et al. 2010; Vandersteen et al. 2010), and assessing habitat requirements for population persistence (Crandall et al. 2000). These understandings support the idea that identifying loci influenced by local adaptation can guide reintroduction and translocation events of captive raised individuals in efforts to repopulate viable habitat and increase genetic diversity and resiliency of populations (Flesch et al. 2020; Weeks et al. 2011). For example, identifying a suite of loci that are consistent

with signatures of selection such as in Dresser et al. (2018) can direct management decisions to choose source-recipient destinations for translocation events based on environmental variables that seem to be dictating local adaptation and thus avoid fatalities of translocated individuals and outbreeding depression. If candidate adaptive loci can be identified in *A. bishopi* populations, that information can be used to inform translocation and reintroduction of individuals to avoid outbreeding depression and repopulate viable habitat.

In contrast to outbreeding, an increase in inbreeding can lead to inbreeding depression and a resulting rise in the risk of extinction due to decreased adaptive potential. Populations are made more susceptible to loss of genetic material via drift and inbreeding through decreased population sizes and increased isolation caused by habitat loss and fragmentation. By identifying *A. bishopi* populations at risk of experience high inbreeding levels, methods may be established to combat inbreeding depression and manage the associated risks to the species' continued survival. I plan to assess small-population risks through estimating population's effective number of breeders (N_b) and genetic differentiation and reconstructing pedigrees. A pond receiving adequate sampling effort and producing relatively large number of sampled larval individuals (40) that produces a small N_b would indicate that few individuals are surviving to adulthood and successfully contributing genetic material to the next generation, which means the population is losing genetic material at a greater rate than a population with a larger N_b. A population with a small N_b should also experience genetic drift more quickly, which would cause higher genetic differentiation between geographically close populations. By identifying ponds at higher risk of experiencing inbreeding depression and local extinction, management actions such as improving habitat quality to increase growth and survival of juveniles and adults, expanding habitat to boost

local population size, or headstarting can be directed in order to combat these risks and preserve these populations.

Landscape Genomics

Landscape genomics is a multidisciplinary field that aims to combine landscape ecology, population genomics, and spatial statistics to understand the effect of environmental change on evolutionary processes. Landscape genomics differs from traditional population genomics in its use of spatially explicit analyses and more sophisticated modeling of spatiotemporal environmental processes as well as its approach of often analyzing data on the individual scale rather than on the population level (Manel et al., 2003; Storfer et al., 2007). Additionally, landscape genomics places emphasis on the role of environmental patterns and processes as driving forces for microevolution (Balkenhol et al., 2015; Balkenhol et al., 2017). In combining landscape ecology and population genomics, landscape genomic studies require landscape data through the form of remotely sensed data, digital landscape models, or field collections as well as genomic data in the form of multilocus genome markers. One of the main goals of landscape genomics is to estimate landscape influences on functional connectivity using gene flow indices such as genetic differentiation and migration (Murphy & Evans 2011).

By applying landscape genomic approaches, we can gain insight on individual dispersal, population fragmentation, and functional connectivity. Through a landscape genomic lens, gene flow is understood by the effects of not only geographic distance but also fine-scale environmental variables such as altitude, topography, and ground cover (Murphy et al., 2010; Funk et al., 2016; Murphy et al., 2018). Geographic distance and environmental data are used in combination with information on how these factors influence habitat fragmentation, patch size, and landscape permeability to gain a wholistic understanding of gene flow in the species and

population of interest. As a result, genetic patterns of diversity are viewed as a response to habitat type, amount, and configuration (Dileo and Wagner, 2016; Balkenhol et al., 2015). The factors affecting population structure such as migration and dispersal can be altered due to varying levels of landscape resistance to gene flow (Zeller, McGarigal, & Whitely, 2012). Varying levels of landscape resistances leads to some areas between populations being more permeable to species movement and therefore allow increased gene flow between populations while other areas are less permeable and decrease gene flow. Investigating landscape resistance helps us to understand geneflow of species with life history traits and movement strategies that are difficult to track in field studies such as fossorial, small-bodied, moisture-dependent species like *A. bishopi*.

Understanding the influence of landscape variables on genetic patterns aids in identifying barriers preventing or reducing gene flow as well as areas that act as corridors and promote gene flow. Barriers to gene flow may be in the form of roads, mountain ridges (Riley et al, 2005; Funk et al, 2005), and microhabitats that are impermeable to movement because they exceed tolerance thresholds in categories such as moisture or temperature (Watts et al., 2015). Identifying these barriers to gene flow has implications for conservation, ecological, and evolutionary studies in understanding species movement and management (Dodd et al, 2004; Walker et al., 2003; Kreyer et al., 2004; Funk et al., 2005). For example, Watts et al. (2015) found genetic connectivity between boreal chorus frog (*Pseudacris maculata*) breeding wetlands to be related to landscape moisture and topographic roughness between wetlands suggesting the importance of snowmelt to gene flow between wetlands. These breeding wetlands seem to follow a stepping-stone connectivity model which highlights the necessity of retaining wetland habitat between occupied ponds to maintain genetic connectivity between pond populations. Identifying

environmental variables such as these that contribute to genetic connectivity of *A. bishopi* breeding wetlands could direct landscape management to preserve conditions pertinent to gene flow and survival of the species.

Beyond understanding dispersal and gene flow using neutral genetic markers, landscape genomic approaches also can identify the action of natural selection through genotype-environment association (GEA) analysis. This method is used to identify candidate adaptive loci which covary with environmental factors (Rellstab et al., 2015; Lasky et al., 2015). Using a GEA approach to landscape genomic analysis is often more powerful than differentiation-based outlier detection in identifying adaptive loci (De Mita et al., 2013; Forester, Lasky, Wagner, & Urban, 2017). Identifying loci associated with environmental factors can improve management in 1) conserving evolutionarily important genes in the population, 2) directing attention to specific habitat variables contributing to environmental gradients, and 3) determining donor-recipient population pairings for species translocations (Harrisson et al., 2014). Uncovering adaptive information useful to management decisions can be exampled by the GEA analysis results in Harrisson et al. (2017) which identified genome regions associated with temperature, precipitation, and geography in Murray cod (*Maccullochella peelii*) across its geographic range, suggesting that these environmental conditions are particularly important to consider when planning translocation events. Similarly, applying GEA techniques to *A. bishopi* populations could uncover pond-level environmental conditions that are influencing local adaptation. Such information would be crucial to conservation actions to preserve adaptive potential and improve evolutionary resilience of the species as well as avoid outbreeding with translocation decisions.

Considering the dual effects of landscape structure on both genetic diversity and differentiation is an often-over-looked use of landscape genomics. Landscape genomic studies

have provided useful information but gaps in its application remain. Many landscape genomic studies look at the influence of landscape permeability on genetic differentiation between populations (Balkenhol et al., 2015). These studies often use a link-level analysis to see if the intervening landscape effects populations further than separating a panmictic population into discrete units and typically look at genetic differentiation (F_{ST}) correlations with landscape variables (Wagner and Fortin, 2013). However, few studies try to understand the effects of habitat amount and configuration on genetic diversity within populations such as expected and observed heterozygosity (H_E , H_O), number of alleles (A) and allelic richness (A_R) (Manel and Holderegger, 2013; Pflüger and Balkenhol, 2014; Keon, Bowman, and Wilson, 2015; Dileo and Wagner 2016). Understanding the effect of landscape composition on population genetic diversity is important since these estimates typically hold our baseline understanding of individual and population level fitness, extinction risk, and ability to respond to change (Frankham, 2005; Reed and Frankham, 2003). Thus, by obtaining information on how and to what degree the landscape influences these genetic estimates, we also deepen our understanding of population status. For example, although Wendt et al. (2021) investigated the influence of pond isolation and suitable-habitat-amount on H_E and A in populations of *A. bishopi* on Eglin. I aim to increase our understanding by expanding the extent of the analysis to include a broader range of populations, including a larger number of genomic markers, and applying newer methods of landscape analysis that optimize landscape resistance values using pairwise genetic data and random-walk commute times without *a priori* resistance assumptions (Peterman, 2018).

Conclusion

The overarching goals of this thesis are to use population, conservation, and landscape genomic techniques to a) enhance our range-wide understanding of genomic relationships within

and between populations of *A. bishopi* and b) evaluate in detail the metapopulation dynamics, genetic diversity, demographics, and conservation status of the species in two focal geographic areas (Eglin and Escribano). Ultimately, this study aims to provide guidance on how to conserve this species and its habitat more effectively. These overarching goals are approached through utilizing genomic techniques coupled with extensive and intensive sampling of *A. bishopi* populations to 1) describe population genetic structure across multiple scales, 2) estimate population genetic diversity and demographic indices, 3) evaluate the efficacy of headstarting programs, 4) assess between- and within-breeding wetland habitat conditions as they relate to genetic connectivity, and 5) as "final frontier" of this study, identify genome regions under natural selection. The results of these components answer evolutionarily significant questions to fill *A. bishopi* knowledge gaps and inform management decisions regarding this vulnerable species. In pursuit of these goals, I was also able to design a SNP panel for target sequencing via RADcap protocol as well as evaluate the application of RADcap for high-throughput sequencing and analysis of a particularly large-genome study organism.

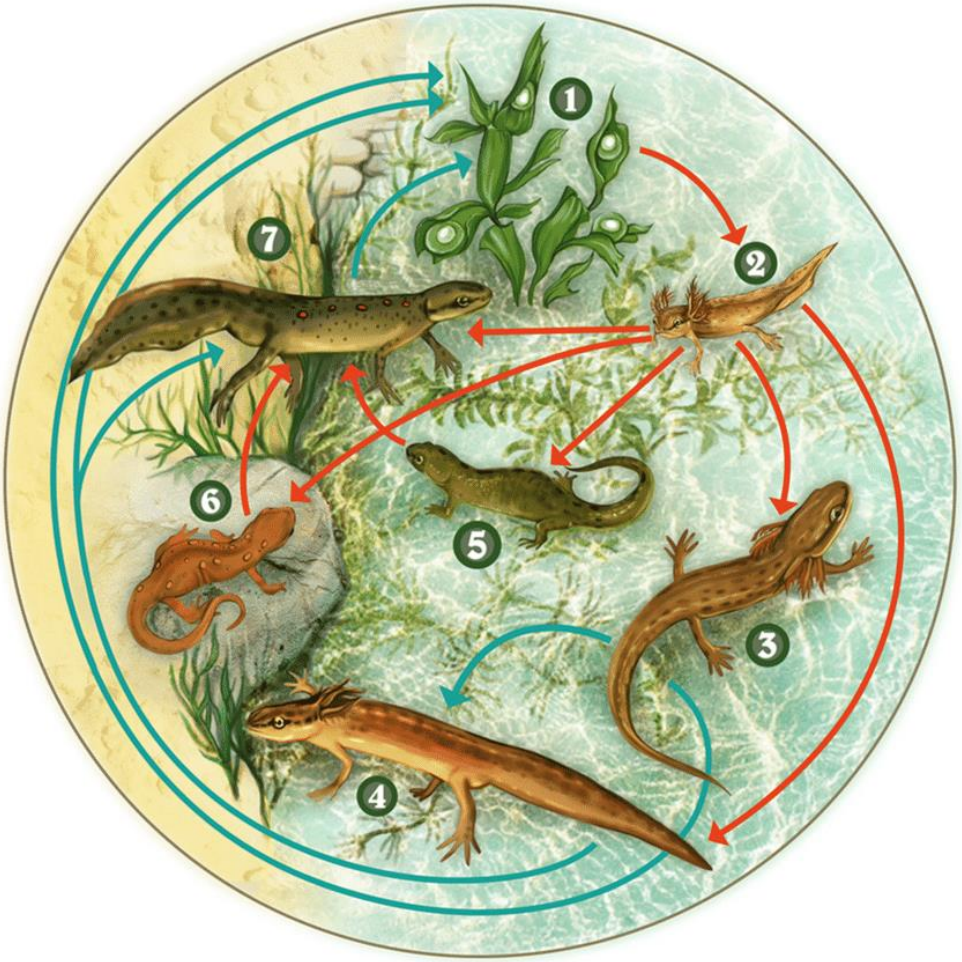

Figure 1.1. Taken from Bohenek and Resetarits (2018). Typical life cycle of an Eastern Newt (*Notophthalmus viridescens*). 1) Eggs are laid in aquatic vegetation. Eggs hatch into 2) aquatic larvae. There are five phenotypes for post larval development. Not all amphibians have life cycles that are this complex. However, the five post-larval phenotypes of the eastern newt depicts many of the common post-larval phenotypes of other amphibian species. The four phenotypes are 3) paedomorphs with full larval morphology with gill slits, external gills, and tail fin; 4) paedomorphs with no gill slits, partially absorbed gills and tail fin; 5) metamorphosed, aquatic juvenile with small, compressed tailfins; 6) metamorphosed, terrestrial efts with dry, hydrophobic skin and 7) metamorphosed, semi-aquatic adults. Only semi-aquatic adults are sexually mature and can reproduce to lay eggs.

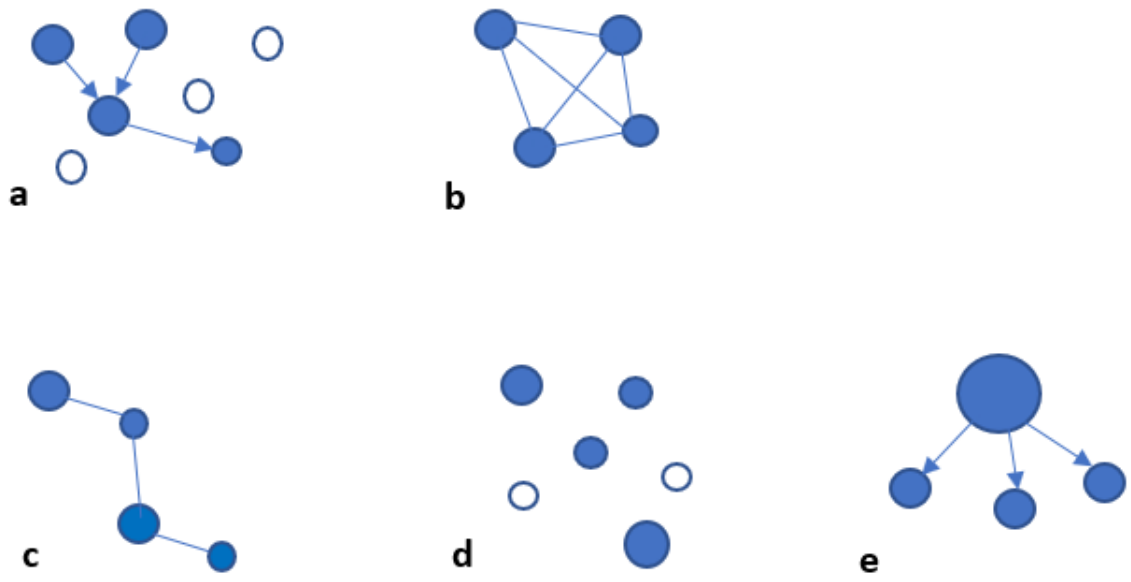

Figure 1.2. Adapted from Harrison (1991). Depictions of metapopulation dynamics. Closed circles represent habitat patches: filled=occupied, unfilled=vacant. Lines represent migration/dispersal: arrowheads= asymmetrical migration, no arrowheads=symmetrical migration. Many combinations of these models are possible. *a)* Levins style metapopulation *b)* panmictic/patchy population *c)* isolation by distance metapopulation *d)* isolated/nonequilibrium populations *e)* source-sink/mainland-island.

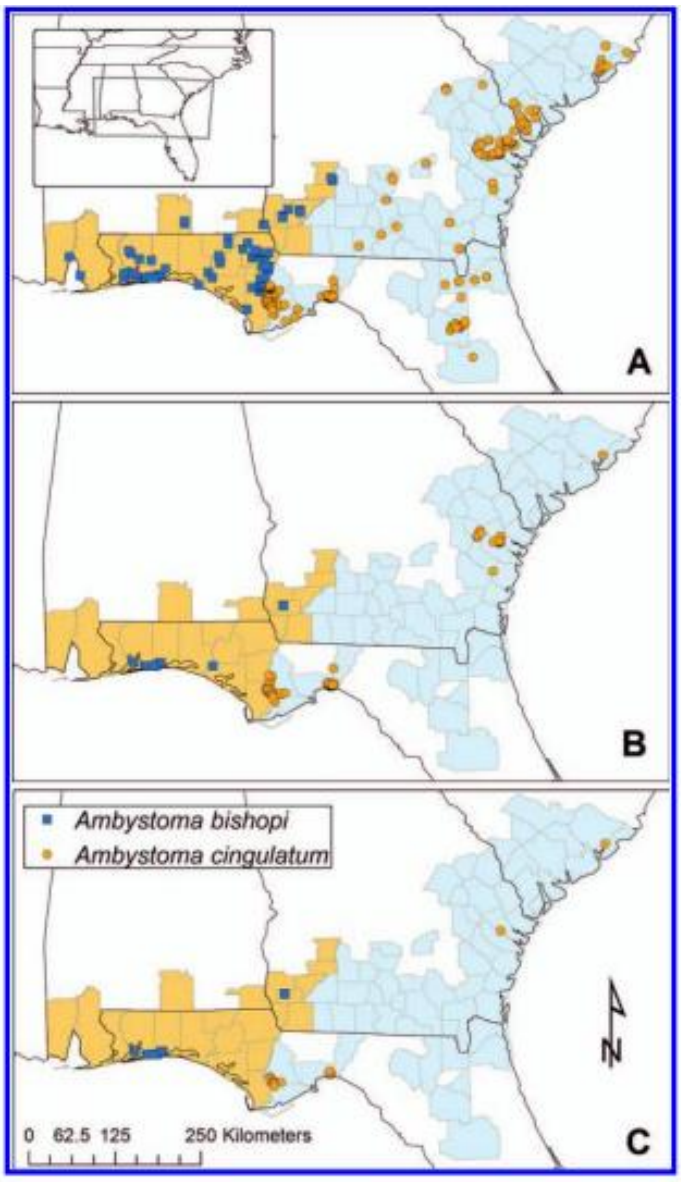

Figure 1.3. Taken from Semlitsch et al. (2017). The known localities of *Ambystoma cingulatum* and *Ambystoma bishopi* over three time periods. (A) all known records (B) 2000-2009 (C) 2010 to 2015. Orange circles indicate *A. cingulatum* records and blue squares indicate *A. bishopi* records. Polygons indicate county lines. Shaded counties indicate the species' range.

CHAPTER 2

DEVELOPMENT AND EVALUATION OF A TARGETED SNP PANEL FOR AMBYSTOMA BISHOPI

Introduction

The reticulated flatwoods salamander (*Ambystoma bishopi*) is federally listed as endangered as a result of its declining range due to habitat loss, fragmentation, and alteration (Pauly et al. 2007; Semlitsch et al. 2017). The last remaining strongholds of this species are found on Eglin Air Force Base (Eglin) and Escribano Point Wildlife Management Area (Escribano), both located in Northwest Florida. Previous genetic research on this species has used microsatellite and mitochondrial markers to delineate population boundaries, assess genetic diversity and viability, and estimate population size (Wendt et al., 2021; Williams et al., 2021). However, findings of these studies may be improved through using high-throughput genomic sequencing techniques in order to obtain thousands of both neutral and adaptive genomic markers across thousands of individuals throughout the species extant range.

Salamanders have some of the largest tetrapod genomes, ranging from 9.9 gigabases (Gb) to 118 Gb (www.genomesize.com), whereas most mammals have 3-4 Gb genomes (Kapusta et al., 2017; Gregory 2018). The size of these genomes is mostly attributed to the large number of long terminal repeat retrotransposons (Sun et al 2012; Sun and Mueller 2014) and long introns (Smith et al 2009; Nowoshilow et al 2018). We can presume *A. bishopi* has a genome close to ~30 Gb based on the *Ambystoma mexicanum* genome, which is 32 Gb (Nowoshilow et al. 2018).

A. bishopi's large genome presents challenges in DNA sequencing and analysis. Restriction-site-associated DNA sequencing (RADseq) and sequence capture are two separate DNA sequencing methods used to sequence reduced amounts of the genome. However, there are key limitations to using these methods in isolation to sequence large-genome species. In RADseq

methods, there are many more potential restriction enzyme cut sites in a larger sized genome, which produces more DNA fragments to sequence. With a greater number of fragments, the number of individuals that can be sequenced per sequencing lane while maintaining adequate coverage per RAD locus of interest is reduced. Furthermore, with hundreds of thousands of sequenced loci, computation and analysis time is slowed. Sequence capture reduces the number of sequenced loci by using adapter-like probes to target specific loci to retain for sequencing. However, sequence capture in a large genome causes capture probes to be diluted, leading to a high amount of "off-target" captures which causing a low yield of target loci enrichment and insufficient depth of read coverage (Guo et al. 2012; Weisrock et al., 2018). Because of these reasons, the application of either of these methods by themselves is not ideal for high-throughput sequencing of a large-genome species.

The challenges presented by the large genome *of A. bishopi* may be best matched through combining RADseq and sequence capture methods and using taxon specific capture probes. The combination of these sequencing methods is termed RADcap (restriction-site-associated DNA capture; Hoffberg et al., 2016). Through applying RADcap methods, target loci can be reliably sequenced across many more individuals at once than with RADseq or sequence capture alone, while still producing deep sequencing coverage of each individual (Hoffberg et al., 2016; Ali et al., 2016; Komoroske et al., 2019).

A typical genome-scale study will uncover a combination of loci that appear to be under selection and loci that appear to be selectively neutral; each type is useful for answering different questions. Neutral markers are useful for questions related to genetic drift and gene flow including migration patterns, dispersal, effective population size, inbreeding levels, neutral population genetic structure, neutral genetic diversity, and neutral genetic differentiation (Slate,

2004, Luikart et al., 2018). Adaptive markers are useful for answering questions related to local adaptation, genetic differentiation caused by adaptation, and the diversity of adaptive alleles. Genomic regions affected by local adaptation are being selected due to local environmental conditions. These adaptive regions can drive gene flow and genetic differentiation within a metapopulation which would not be detected when analyzing neutral loci alone (Hoffmann et al., 2003; Chapman et al., 2009). By creating sequence capture probes specific to *A. bishopi* and targeting both neutral and candidate adaptive loci, a more thorough and complete understanding of *A. bishopi* genomic status and underlying mechanisms driving diversity and gene flow can be obtained. The purpose of this chapter was to 1) identify putatively neutral and adaptive loci based on RADseq of individuals from all extant populations, 2) design a targeted RADcap panel based on these loci and evaluate their performance. This panel was then used in high-throughput sequencing to genotype hundreds of additional individuals to address objectives posed in Chapters 3 and 4.

General Methods

Study Area

This study analyzes individuals from all known extant major population centers (Mayhaw, Garcon, Eglin, and Escribano) to answer questions regarding range-wide diversity and relatedness of *A. bishopi* populations. The relative locations of these regions can be viewed in Figure 2.1. However, populations occupying Eglin and Escribano ultimately were analyzed more intensively to answer questions at the pond and metapopulation levels due to the number of active breeding ponds on these properties (Chapters 3 and 4).

The populations of *A. bishopi* on Mayhaw and Garcon are considerably smaller than the populations on Eglin and Escribano. The *A. bishopi* critical habitat unit in Mayhaw is a circular

66 ha area delineated around a single historic breeding wetland and is managed by the Georgia Department of Natural Resources. The most recent detection of *A. bishopi* at Mayhaw included individuals from two separate wetlands populations (John Jensen, GADNR, pers. comm. 2017). Prescribed burns in this area help to provide continued suitable habitat for *A. bishopi* on this property (Means, 2013). Garcon contains one breeding wetland where *A. bishopi* has been found and another wetland with no records of *A. bishopi* presence. Samples from this location consisted of one sampling event from a single pond. Nine additional wetlands suitable for *A. bishopi* breeding exist on this property identified by Palis and Enge (2006). The populations of *A. bishopi* at these properties may be considerably smaller than Eglin and Escribano populations, however Williams et al. (2020) found unique MHC (major histocompatibility compex) alleles in the Mayhaw and Garcon Point populations, highlighting the importance of prioritizing these breeding sites for conservation in order to retain MHC diversity in the species.

Eglin is the largest forested Air Force installation in the United States, with 464,000 acres of land and 120,000 square miles of water (U.S. Department of the Air Force, 1998). Eglin contains the largest contiguous area of old-growth longleaf pine in the United States and is home to over 106 rare and federally threatened and endangered plant and animal species (Hiers et al., 2003; Florida Natural Areas Inventory). There are two main metapopulations of breeding ponds on Eglin: Oglesby and Eastbay. Wendt et al. (2021) found these two areas to be genetically distinct metapopulations with no contemporary gene flow between the two regions. However, there is another aggregation of ponds on an Air Force installation adjacent to Eglin collectively referred to as “Hurlburt Field”. The relationship of these ponds to the Oglesby and Eastbay metapopulations has not been studied and may add to our understanding of gene flow on the property. Oglesby contains 8 ponds that are known to have been used for breeding in the past 5

years, (ponds 4, 5, 49, 51, 52, 53, 212, 213). Eastbay contains 9 ponds that are known to have been used for breeding in the past 5 years (ponds 15, 16, 19, 32, 33, 34, 112, 215, 234). Over 10 km of unsuitable habitat separates Oglesby and Eastbay metapopulations. The four Hurlburt ponds in this study (H4, H5, H6, H8) are 5.9 km away from Oglesby and 4.0 km from Eastbay.

The Oglesby and Eastbay metapopulations on Eglin have received by far the greatest ecological and population genetic study. Hydrological conditions have been measured at ponds to determine which ponds have longer hydroperiods and recession rates and how this may affect breeding pond suitability (Chandler et al., 2017). Mechanical woody vegetation removal has been compared to fire-treatment in their effectiveness in restoring amphibian breeding habitat (Gorman, Haas, & Himes 2013). Brooks et al. (2019) investigated the influence of environmental variables on *A. bishopi* occupancy and apparent dispersal. Other studies include investigating movement and burrow use (Powell, Gorman, & Haas 2015), assessing feral swine presence and associated damage to breeding ponds (Jones et al., 2018), descriptions of egg deposition sites (Gorman et al., 2014), and population genetic investigations using microsatellite and immune-related markers (Wendt et al., 2021; Williams et al. 2020).

Escribano also appears to harbor a large aggregation of active breeding sites but historically has received less study or management. However, Williams et al. (2020) included samples from Escribano and other areas to investigate immune gene presence and adaptation to pathogens. This study found low levels of diversity compared to other amphibian species at MHC exons, which suggests that the species may have an elevated infectious disease risk (Savage et al. 2019). Escribano consists of 4,057 acres along the shorelines of the waterbodies Blackwater Bay and East Bay and contains old growth longleaf pine and slash pine habitat along with salt marsh, dome swamp, shrub bog, mesic flatwoods, scrubby flatwoods, and sandhill

habitat. Escribano contains 15 breeding ponds (EP1, EP5, EP15, EP46, EP47, Gum, Honey Hole, Ghost, Tadpole, Torpedo, Restoration, Banana, Ditch, Borrow, and Stanley) and a substantial number of wetlands that have yet to be sampled.

Managers on both Eglin and Escribano aim to maintain the longleaf pine ecosystem through controlled burns, mechanical removal of woody vegetation, and chemical prevention of woody vegetation regrowth. *A. bishopi* monitoring activities occur at both Eglin and Escribano including dip net, drift fence, and spotlight surveys to observe *A. bishopi* presence and collect tissue samples. Headstarting programs occur with each breeding season at both sites beginning with egg and larvae collection, continuing with captive rearing, and ending with release into source ponds or unoccupied translocation sites at the late metamorph or juvenile phase. Monitoring and headstart activities began at Eglin in 2003 and 2015 respectively and more recently began at Escribano in 2018.

Sample Collection

Tissue samples were collected non-lethally by clipping a small section of the tail, which regenerates. Samples were collected by Escribano and Eglin biologists and U.S. Fish and Wildlife personnel from the 21 active breeding ponds on Eglin (8 on Oglesby, 9 on Eastbay, and 4 on Hurlburt field) and the 15 known, active breeding ponds on Escribano. Relative locations of breeding ponds can be viewed in Figure 2.1. Tissue was stored in 95% ethanol at -20˚ until DNA extraction. Specimens were found using drift fences, dipnet surveys, and flashlight searches. All tissue samples included in this project were collected at late-stage larvae or metamorph stage. Individuals used for the headstarting program were collected as eggs during the breeding season and subsequently reared in cattle tanks. Individuals were separated into cattle tanks by pond of origin and are released into the same pond from which they were initially collected. Each head-

started individual was tissue-sampled before release and their life stage, date sampled, and pond were documented. Ponds used in the headstarting program at Eglin and Escribano can be seen in Figure 2.2. Sampling efforts were performed between October and May each year. Tissue from five sampling years (2016, 2017, 2018, 2019, 2020) are included in this study, though sample sizes per pond and year varied widely due to spatiotemporal variation in recruitment and sampling effort. Because the breeding season of *A. bishopi* begins in late fall and continues into spring, breeding seasons are named for the fall year (i.e., the breeding season spanning Fall 2018 to Spring 2019 is named the 2018 breeding season).

Tissue sampling at Escribano involved an additional step to address a goal of Chapter 3—determining what percentage of the origin population's diversity is represented in headstart samples. As at Eglin, individuals that were captively reared in cattle tanks were collected as egg masses through search efforts conducted early in the breeding season before pond inundation. However, at Escribano, an additional round of sampling occurred later in the breeding season. Once breeding ponds filled with water, dipnet surveys were conducted in order to sample individuals hatched from eggs laid later in the season or were well-hidden and missed by egg searches. These individuals that were sampled later in the breeding season are referred to here as "naturally reared", whereas individuals that had hatched and grown in cattle tanks are referred to as "captively reared". By sampling ponds used for headstarting programs later in the breeding season, I assessed if captive populations are equally diverse as the entire source population.

On Escribano, eight ponds were used in the headstarting program (Torpedo, Honey, Ghost, Borrow, Ditch, Stanley, Gum, and EP15) and seven ponds were excluded from the headstarting program (EP1, EP5, EP46, EP47, Restoration, Banana, and Tadpole). Ponds that were not used in the headstarting program are referred to here as "wild" ponds. Both headstart

ponds and wild ponds were sampled over multiple years. However, only 2019 produced sample sizes from wild ponds adequate for genetic analysis. The variable number of samples from ponds each breeding season is potentially due to several factors including ongoing landscape management leading to the discovery of and access to these wild ponds, the boom-bust breeding nature of pond breeding amphibians, and high mortality rate of eggs and larvae. On Eglin, two ponds were used in a headstarting program with augmentation (similar to the program at Escribano), and larvae were reared in cattle tanks and released back into the source ponds at P4 and P5 starting in 2014 and continuing through spring 2019. Starting in 2019, Eglin began headstarting for translocation into ponds where populations had been extirpated for over 10 years. At Eglin, ten ponds were used to collect eggs and larvae for the headstarting program (P4, P5, P15, P16, P33, P34, P53, P112, P212, P215) and ten ponds were not used in the headstarting program (P19, P32, P51, P52, P213, P234, H4, H5, H6, H8). Early in the 2019 breeding season, individuals from ponds 5, 15, 16, 33, 52, 53, 212, and 215 were transplanted to the inactive breeding ponds 2 and 30 on Eglin. Ponds 2 and 30 were subsequently surveyed for samples later in the breeding season. Thus, all tissue samples collected from ponds 2 and 30 represented first generation migrants. Pond 49 is an active breeding pond that received transplant individuals from pond 5 during the 2016 breeding season but has not received transplant individuals since. For these reasons, ponds 2, 30, and 49 are referred to as transplant ponds.

DNA Extraction

Whole genomic DNA was extracted from tissue samples using DNeasy Blood and Tissue Kits (Qiagen, Hilden, Germany). To increase DNA concentrations, I performed two elution steps for each sample in attempts to increase DNA quantity with the additional pass through the DNeasy membranes (DNeasy Blood & Tissue Handbook) and I eluted DNA into water to allow

for concentration via evaporation using a Savant SpeedVac SC110. Even so, many tissue samples were small and produced low concentrations of DNA (0.1- 5 ng/uL).

Designing sequence capture baits

Pilot Library Preparation

In order to design RADcap baits, I first used 3RAD, a typed of dual-digest RADseq technique (first used by Hoffberg et al. 2016 and developed by Travis Glenn of UGA http://baddna.uga.edu/faq.html) to identify candidate RAD loci suitable for use in a targeted SNP panel. Individuals were chosen for bait design with the goal of representing each of the four properties and multiple ponds from Eglin and Escribano across multiple years in order to minimize geographical and temporal bias in selecting target RAD loci. In all, 55 *A. bishopi* individuals were subjected to 3RAD, across two libraries. Triple-digest-restriction-site-associated DNA protocol creates a double-digest Restriction Associated DNA sequence library by using two digest enzymes and then utilizes a third enzyme to cut primer-dimers and increase efficiency (Bayona-Vásquez et al., 2019). For this study, I used enzymes Xba1, EcoR1, and Nhe1 accompanied by adapters iTru Nhe1, iTru EcoR1, and iTru7 as well as a random 8 nucleotide index. The standard 3RAD protocol was followed, except for an additional two PCR cycles in the amplification step to ensure sufficient final DNA concentration (Bayona-Vásquez et al., 2019). Pippin prep (Sage Science, Beverly, Massachusetts) was used to size-select for 525-bp fragments, followed by Qubit (Thermo Fisher Scientific, Waltham, Massachusetts) to quantify concentration and Bioanalyzer (Agilent Technologies, Santa Clara, California) to check fragment size before sending to the University of Oregon Genomics Core facility for Illumina HiSeq 4000 150-bp paired-end sequencing.

De novo assembly and SNP calling

After sequencing, Stacks version 2.3e (Catchen et al., 2013) was used to process raw genomic data. Data were demultiplexed into individual samples and PCR duplicates were filtered followed by the Stacks de novo pipeline. In total, 774,882,218 raw reads were demultiplexed into individual samples using barcode sequences. Of these, 690,260,279 reads were retained (89.1%) for decloning and used as input for the Stacks de novo pipeline. The Stacks de novo pipeline aligns reads to identify loci containing SNPs and removes poor quality reads through applying a number of filters that SNPs must pass through to be included in the final dataset. This is done through several phases: Ustacks, Cstacks, Sstacks, Gstacks, and Populations which can be seen in Figure 2.3.

Based on 3RAD library synthesis and sequencing of these initial 55 individuals, I optimized settings for the following methods and conducted preliminary analyses to develop the RADcap panel. Stacks parameter optimization was implemented by testing a range of values for stacking filters (m, M, N, n, and R). I tested m=2-5, M=1-8, N=4-7, n=2-5, and R=80-90. The role of these filters can be viewed in Table 2.1. After trying different combinations of Stacks parameters, M=3 and R=90 provided the greatest number of SNPs while filtering out low-quality SNPs (SNPs with low depth of coverage) and those not found in <90% of individuals. Additionally, reads were trimmed to 140 bp to be uniform size and remove base pairs prone to error in Illumina datasets. For all other parameters in Ustacks, Cstacks, Sstacks, and Gstacks, the default settings were used (appendix table 2.1). As a result of Ustacks, Cstacks, Sstacks, and Gstacks, 782,453 loci were assembled consisting of 920,806,586 paired-end reads with an average 1,173.4 reads per locus and mean effective per-sample coverage of 32.3x. Using the Stacks populations program, I retained only SNPs with a minor allele frequency ≥0.05 and a heterozygosity ≤0.5 and I retained only the first SNP observed per RAD locus to minimize

linkage in the dataset. After the Stacks pipeline and populations filtering, 53,646 SNP markers, each located on a separate RAD locus, were retained for further consideration. Because each locus harbored only one retained SNP, for the remainder of this thesis, I use "locus" and "SNP" interchangeably.

Identification of loci to target for RADcap

In order to identify and remove library effects (i.e., between-library variation in genotype calls due to sequencing errors, PCR variation, enzyme efficiency, read-depth variation, etc.; Bonin et al., 2004; O'Leary et al., 2018), 10 individuals were replicated across the two 3RAD libraries. These replicate individuals were analyzed for genotyping mismatches using custom scripts in R. Examination of redundant genotypes allowed me to account for potential variation between libraries and sequencing runs by utilizing bioinformatic tools to remove erroneous genetic structure caused by the stochastic errors inherent to different sequencing runs and library preparations (O'Leary et al., 2018; Pfeiffer et al., 2018; Puritz et al., 2015). I identified 1,738 SNPs which produced conflicting genotypes between duplicate individuals more than 10% of the time. These SNPs were then removed from consideration for the RADcap panel. Potential paralogs were identified and removed from the dataset using the *HDplot* package (McKinney et al. 2017) in R and a +/- 5 threshold, which removed 1,094 loci. Individuals were grouped by property and each locus was tested for Hardy-Weinberg equilibrium at the 0.05 significance level using Arlequin version 3.5 (Excoffier, Laval, and Schneider 2005) which removed 311 loci. *PCAdapt* (Luu, Bazin, and Blum 2017) was used to test for outlier loci based off of PCA structuring and BayeScan version 2.1 (Foll & Gaggiotti 2008) was used to test for outlier loci using a Bayesian approach. For both these methods, a false discovery rate of 0.1 was used. These methods identified 5,128 and 72 loci as outliers respectively. Due to the large number of SNPs in

the dataset and to be conservative, a locus was removed from the neutral-SNP dataset if it was identified as an outlier in any of these analyses or in the RDA analysis described in the next section. Rarefaction analyses were conducted in order to determine the number of neutral SNPs required to accurately estimate expected heterozygosity (H_E) and genetic differentiation (F_{ST}). The results of these analyses were used to determine how many neutral loci would be targeted using RADcap baits.

In order to uncover correlation between loci and environmental variables and identify potentially adaptive SNPs, a redundancy analysis (RDA) was conducted in the R package *vegan* (Oksanen et al. 2019) following the methods of Forester et al. (2016). This analysis was based only on Eglin due to the absence of relevant environmental data from Escribano. If SNPs were correlated with environmental variables at Eglin, they were assumed to have the same relationship in Escribano individuals. RDA was run using environmental variables of pond recession rate, pond hydroperiod, pond suitable habitat area, fish abundance, and invertebrate abundance. These data were collected through multiple studies on Eglin's *A. bishopi* population including Chandler et al. (2017) and Brooks et al. (2019). Recession rate and hydroperiod affect timing of metamorphosis and duration of the larval stage, whereas fish and invertebrate abundance represent predation levels. Pond area may have an effect on the relationship among larval density, prey items, and habitat resources. Candidate outlier SNPs were identified as those loading much more strongly with environmental variables (i.e. RDA axes) than did the majority of SNPs (most of which should be neutral, and thus more influenced by space). SNPs with a loading >3 standard deviations from the mean were considered outliers. This identified 494 loci to be correlated with at least 1 of the 5 environmental variables.

In consultation with scientists from Arbor Biosciences (Ann Arbor, Michigan) I designed a panel of 6,000 targeted SNP loci, including 5,748 putatively neutral loci and 252 candidate adaptive loci. Neutral loci had passed all outlier tests (library consistency, *HDplot*, *PCAdapt*, Bayescan, and RDA) leaving 45,519 SNPs to select from. Candidate adaptive loci included the 494 SNPs identified through RDA. Arbor conducted analyses to further filter out loci exhibiting an unfavorable GC content or melting point or that matched multiple positions on the *Axolotl* genome. The final SNP panel consisted of the 252 remaining candidate adaptive loci plus a random subsample of 5,748 neutral loci that passed these filters. Once loci were chosen, Arbor Biosciences manufactured the custom bait kit.

Bait performance assessment

Laboratory Protocol

To test these sequence capture baits, RADcap was performed on a library of 264 samples (245 unique individuals, 19 individuals duplicated from the earlier 3RAD library) using the 3RAD protocol (Hoffberg et al., 2016) followed by bait application according to the myBaits® Hybridization Capture for Targeted NGS Manual v.5.0 standard protocol. A hybridization temperature of 65˚C was used for ~23 hours. Following capture, P5 and P7 primers and KAPA HiFi HotStart ReadyMix were used to amplify bead-bound loci in 14 cycles of PCR. PCR products were cleaned with AMpure SpeedBeads. This library was sent to the University of Oregon Genomics Core facility for Illumina HiSeq 4000 150 bp paired-end sequencing.

Bioinformatics

Sequence reads were processed and analyzed using a custom pipeline implementing Stacks, BWA (Li and Durbin, 2009), samtools (Li et al., 2009), VCFtools (Danecek et al., 2011), IGV (Thorvaldsdóttir et al., 2013), and R. A flow chart of this pipeline can be seen in Figure 2.4.

Stacks was used to demultiplex data into individuals, remove PCR duplicates, and trim reads to 140 bp. A FASTA-formatted file was created according to that in Hoffberg et al., (2016) using the 6,000 paired sequences used for bait design as a "reference genome" (referred to as the reference from here on) in read alignment and SNP calling. In order to compare RADcap to 3RAD, reads for both RADcap and 3RAD individuals were aligned to the reference using the BWA-MEM algorithm (Li, 2013). Alignments were then organized using samtools to mimic the positional order of the reference. Alignments for four individuals were indexed and viewed in IGV in order to visually assess for paralogous loci. A locus was labeled as paralogous if it contained a high density of variants. Out of the first 100 loci I examined, 11 loci appeared to be paralogs (11%). However, only 3 of these 11 loci contained SNPs that were actually called following the rest of the pipeline (i.e., the genotype caller usually correctly discarded paralogous variants). Therefore, I assumed that the proportion of paralogous loci in the dataset was no more than 3%. Sorted alignments were fed into the Stacks reference pipeline to assemble loci and call SNPs. Individuals missing ≥30% of SNPs and loci with low coverage (≤10x) were identified using VCFtools and removed. Individuals from the Mayhaw and Garcon sites were also removed as and they were not needed for the 3RAD/RADcap comparison and their inclusion could unnecessarily bias further filtering. Applying the Stacks Populations module, SNPs were then retained if they 1) were present in at least 70% of individuals in both the 3RAD and RADcap datasets, 2) had a minor allele frequency ≥0.05, and 3) had a maximum observed heterozygosity of ≤0.5. For this assessment, candidate adaptive loci were then removed using VCFtools in order to compare only neutral SNPs. H_O and H_E were estimated using *hierfstat* (Goudet, 2005) in R separately by pond and method (3RAD vs RADcap). Individuals duplicated between 3RAD and RADcap libraries were then isolated for further analysis. Custom R code was used to calculate

the rate of mismatched genotypes between methods and identify mismatched SNPs. Out of the first 300 loci, 17 genotypes were mismatched between 3RAD and RADcap methods (5.7%). Out of those 17 mismatched genotypes, only 3 were "true" mismatches (heterozygous in one method and homozygous in another method; 17.6%) while the other 14 were due to missing RADcap data. Loci were then removed from the dataset if they were missing from ≥5% of individuals. H_O was then estimated for duplicate individuals using *hierfstat* and a PCA was created to evaluate genomic variation between individuals.

High-throughput library preparation

RADcap was performed to collect genomic information on the remaining individuals. In total, nine RADcap libraries were created for this project, each consisting of between 192 and 288 individuals. After the initial RADcap library, subsequent libraires were sequenced on an Illumina NovaSeq 6000 to produce 150-bp paired-end reads, as HiSeq sequencing was no longer offered at the University of Oregon's Genomics Core facility after June 2021. Each library shared at least ten duplicate individuals with at least one other library, in order to account for the library effect (O'Leary et al., 2018; Pfeiffer et al., 2018; Puritz et al., 2015). As much as possible, properties, ponds, and years were spread across libraries and plates in order to minimize the total influence of plate and library variance on analysis. In total, 2,255 unique individuals were sequenced for this project. The number of individuals sequenced per pond, per year for each property can be viewed in Table 2.3. Based on the findings of Nazareno et al. (2017) and Nunziata and Weisrock (2018), I targeted 40 individuals per pond per cohort if possible, in order to have adequate sample size for estimating of F_{ST} and heterozygosity, effective populations size, and conducting landscape analyses. For some ponds of particular interest (Borrow, Ditch, Honey, Ghost, EP15, and Gum) I comprehensively sequenced all

captured individuals, to increase accuracy of relatedness estimates, pedigree analysis, and objectives relating to the headstarting program.

Bioinformatics

The bioinformatics pipeline used on the final dataset mostly follows that used in evaluating RADcap as outlined in Figure 2.4, with the following alterations. When applying the Stacks reference pipeline on alignments, the populations program was used to broadly filter the dataset to retain SNPs present in at least 50% of individuals in every library (r=0.5, p=9) with a minor allele frequency ≥0.05 and a maximum observed heterozygosity of ≤0.5 . VCFtools was then used to "whitelist" the targeted SNPs and identify individuals missing ≥30% of SNPs to remove those individuals from the dataset. Loci influenced by library effect were identified by comparing genotypes for individuals replicated between libraries and removing any locus exhibiting a mismatched genotype in more than 20% of these individuals. I then removed one of the two replicated individuals in order to avoid bias in further analysis. The Stacks populations program was then run again using a more stringent filtering approach to retain SNPs present in 80% of individuals across the entire dataset (R=0.8) with minor allele frequency ≥0.05 and maximum observed heterozygosity of ≤0.5. Targeted neutral and candidate adaptive loci were separated using VCFtools.

Results

RADcap bait design

In rarefaction analyses there was relatively little improvement in precision (measured by confidence interval width) beyond 4,000 neutral loci, when estimating H_E (CI decreased by 0.0036 between 4,000 loci and 10,000 loci) and F_{ST} (CI decreased by 0.0033; Figures 2.5). Based off of this finding, I determined 6,000 loci to be an adequate number of total loci (neutral and

adaptive) to target using RADcap baits, assuming that not all targeted loci will be successfully genotyped for every individual.

RADcap evaluation

RADcap pilot library sequencing returned 6074 loci post-alignment and Gstacks, derived from 86,010,652 paired end reads with a mean per sample coverage of 62.9x. After applying minor allele and maximum heterozygosity filters in Stacks Populations, whitelisting targeted loci, and removing individuals with missingness ≥30% (30 individuals), 5,932 targeted loci remained. Of these targeted loci, 223 were candidate adaptive targets leaving 5,709 putatively neutral targeted loci, giving a 98.9% successful return rate of targeted neutral loci and 99.1% for candidate adaptive loci. Returned targeted loci had an average depth of 59.9.

Estimates varied slightly between methods for pond H_O and H_E (Table 2.3) but were almost identical for H_O of all but two individuals duplicated between RAD methods (Table 2.4). Moreover, duplicate individuals closely overlapped in PCA space indicating that genotypic differences would not strongly affect genetic conclusions (Figure 2.6). Due to these findings, I determined RADcap to be an effective and efficient approach for this study.

High-throughput sequencing

Of the 2,255 individuals that were sequenced, 667 individuals did not pass QA/QC standards and were removed due to 1) unexplained whole-library failure, 2) missing barcodes, or 3) high missingness (≥30%). This left 1,588 individuals in the final dataset.

Of the 6,000 targeted loci, 5,127 loci passed filtering parameters applied in Stacks. Of these 5,127 loci, 130 loci were part of the candidate adaptive dataset leaving 4,993 neutral loci. As a result of assessing the data for potential library effects, 141 of these neutral loci were

removed. Additionally, 69 loci were removed because they were not consistently in Hardy-Weinberg equilibrium. As a result, my final neutral dataset consisted of 4,783 loci genotyped across 1,588 individuals. The average individual coverage was 124.0x with 82.9% of individuals having ≥30x coverage and only 57 individuals with coverage <10x. Average site missingness was only 6.6% across all sequenced individuals.

Discussion

Combining RADseq and target capture techniques has been proposed as an adequate sequencing method to address the sequencing challenges presented by large genome species such as *A. bishopi* (Weisrock et al., 2018). Using 3RAD, I was able to obtain tens of thousands of putatively neutral SNPs per individual in order to 1) identify thresholds of importance in the number of loci required to estimate F_{ST} and H_E and 2) design species-specific baits to target putatively neutral SNPs in RADcap. The initial execution of RADcap produced a relatively low number of high missing individuals, high average sample coverage, and a high success rate in returned targeted loci. Heterozygosity estimates on the population level varied slightly between 3RAD and RADcap methods. However, H_O estimates for individuals duplicated between methods was nearly identical for all but two individuals, and duplicate individuals showed little variation in PCA results. The variance in heterozygosity estimates between 3RAD and RADcap methods could be caused by a) the large difference in sample sizes between the two methods, b) an unavoidable allelic imbalance in RADcap data caused by baits favoring one allele over another, or c) the much higher read depth of 3RAD data compared to RADcap data. Due to these findings, I determined RADcap to be an effective and efficient approach for high-throughput sequencing of *A. bishopi*.

In high-throughput sequencing, many individuals had to be removed. Of these individuals, RADcap and sequencing were performed a second time for 254 (one library) in attempts to recover them. However, this did not seem to have an impact on the sequencing quality, and I was not able to recover these individuals. Therefore, I chose to spend this project's limited resources on sequencing other individuals instead of trying to recover individuals that had to be removed. Many individuals (512) had concentrations of DNA <10 ng/μL, which may have influenced their sequencing success. However, not all individuals with low DNA concentrations had to be removed, which suggests DNA quality may have also impacted individuals' sequencing success. All of this highlights the importance of collecting sizable tissue samples, properly storing tissue, and flawless execution of DNA extraction and library preparation. After initial alignment and filtering, there was an 85.5% success rate in returned targeted loci. After more stringent filtering, we saw an 81.9% success rate of targeted loci between candidate adaptive and neutral datasets with very few loci removed due to library effects and Hardy-Weinberg disequilibria. Because of the ability RADcap presents to sequence hundreds of large-genome individuals per sequencing lane while maintaining sequence coverage and the relatively high success rate of returned targeted loci with low site missingness across adequately sequenced individuals, it seems to be an effective and efficient approach for high-throughput sequencing of hundreds of large-genome individuals in order to answer questions of conservation concern.

Table 2.1. Stacks parameters optimized for SNP filtering

Filter	Stacks Step	Setting used	Description
M	ustacks	3	Maximum nucleotide distance allowed between stacks
m	ustacks	3	Minimum depth of coverage required to create a stack
N	ustacks	7	Maximum nucleotide distance allowed to align secondary reads to primary stacks
n	cstacks	1	Number of mismatches allowed between sample loci when building the catalog
R	populations	90, 70	Minimum percentage of individuals across populations required to process a locus
r	populations	50	Minimum percentage of individuals within populations required to process a locus
min-maf	populations	0.05	Minimum minor allele frequency required to process a nucleotide site at a locus (applied to the metapopulation)
Max-obs-het	populations	0.5	Maximum observed heterozygosity required to process a nucleotide site at a locus (applied to the metapopulation)

Table 2.2. Samples sequenced per pond per year for each property.

Property	Pond	2016	2017	2018	2019	2020	Total
Escribano	Borrow			42	46	33	121
	Banana				6		6
	Ditch			70	49	52	171
	Ghost			46	47	152	245
	Gum				79	95	174
	Honey			61	57	55	173
	Restoration				24		24
	Stanley			24			24
	Tadpole				4		4
	Torpedo			39	131	72	242
	EP1				23		23
	EP5				15		15
	EP15				75	1	76
	EP46				36		36
	EP47				20		20
Eglin: East Bay	P15			3	41	40	84
	P16				2	44	46
	P19				2		2
	P30				6		6
	P32				9		9
	P33			23	45		68
	P34				3		3
	P112				10	40	50
	P215		3		46	40	89
	P234				14	1	15
Eglin: Oglesby	P4	42	40	42	8	5	137
	P5	43	40	42	3	1	129
	P49				12		12
	P51				4		4
	P52				28		28
	P53				40	42	82
	P212			3	47		50
	P213				23		23
Eglin: Hurlburt	H4				14		14
	H5				11		11
	H6				8		8
	H8				22		22
Mayhaw							5
Garcon							4
						Total	2255

Table 2.3. Estimates of observed and expected heterozygosity for ponds across years for ponds represented in both 3RAD and RADcap methods. Standard deviations are in parentheses.

Pond	3RAD H_O	RADcap H_O	3RAD H_E	RADcap H_E
Borrow	0.26 (0.24)	0.25 (0.20)	0.27 (0.21)	0.25 (0.20)
Ditch	0.27 (0.26)	0.25 (0.23)	0.28 (0.22)	0.26 (0.20)
Pond33	0.31 (0.23)	0.28 (0.19)	0.32 (0.19)	0.30 (0.17)
Pond4	0.32 (0.23)	0.29 (0.19)	0.33 (0.19)	0.30 (0.17)

Table 2.4. Estimates of observed heterozygosity for individuals that were sequenced in both 3RAD and RADcap methods.

Individual	3RAD H_O	RADcap H_O
Bor-012720-1	0.28	0.28
Bor-012720-3	0.22	0.24
Bor-032719-8	0.15	0.15
Dit-012320-4	0.28	0.22
DIT-032919-7	0.23	0.23
F18-0006	0.31	0.31
F18-0107	0.29	0.29
F18-0345	0.30	0.30
F19-0020	0.32	0.31
F19-0773	0.29	0.28

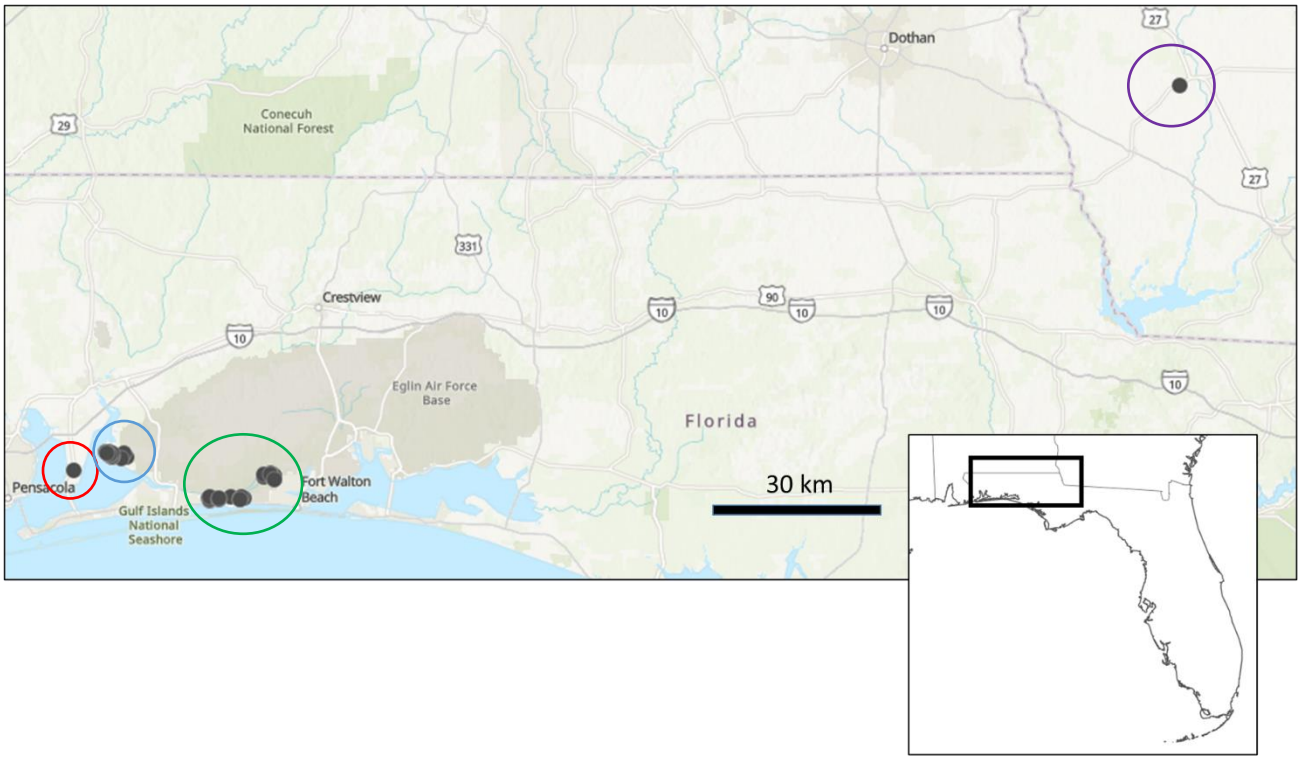

Figure 2.1. Relative locations of each *A. bishopi* population included in this study—Garcon Point (red), Escribano (blue), Eglin (green), and Mayhaw (purple). Garcon Point, Escribano, and Eglin are found in the western section of the Florida panhandle whereas Mayhaw is found in the southwestern corner of Georgia.

Figure 2.2. Sampling sites for Eglin and Escribano. Breeding ponds used for headstarting program on Eglin and Escribano in blue. Breeding ponds not used for headstarting in red.

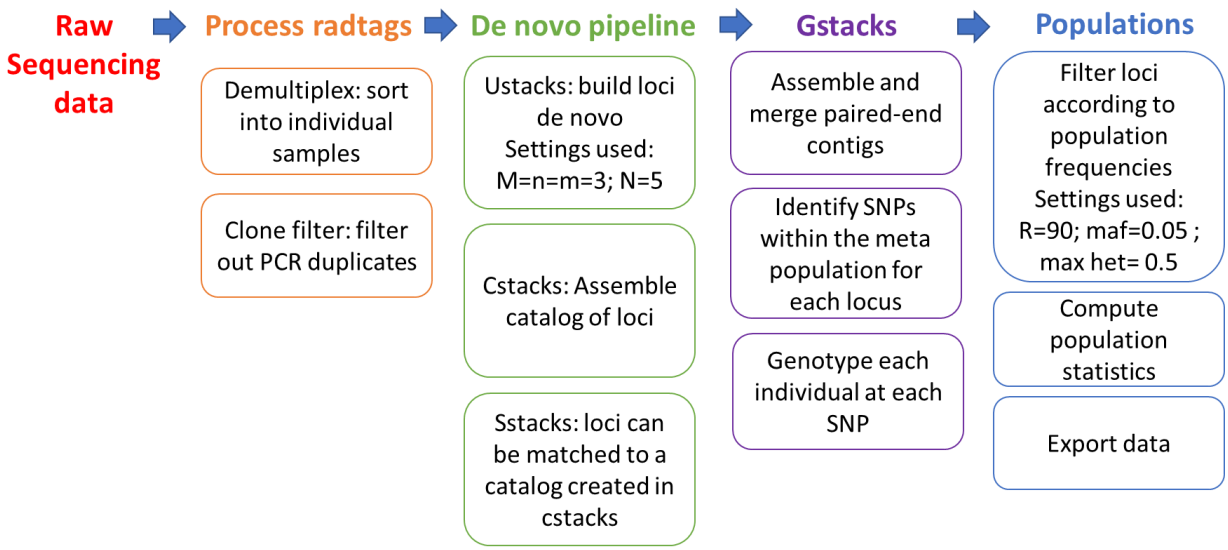

Figure 2.3. Stacks filtering components with key settings used along the pipeline.

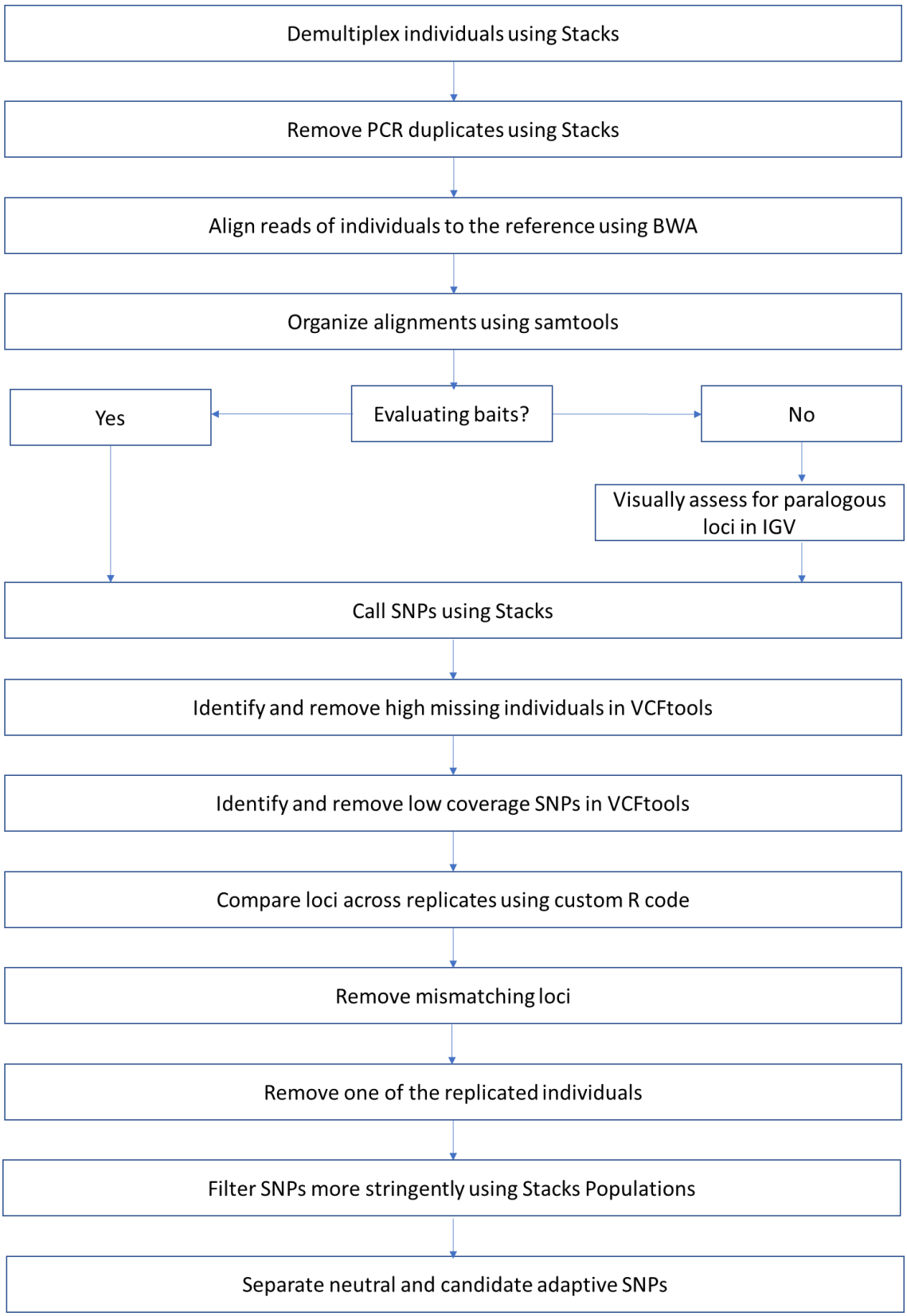

Figure 2.4. General overview of bioinformatic workflow

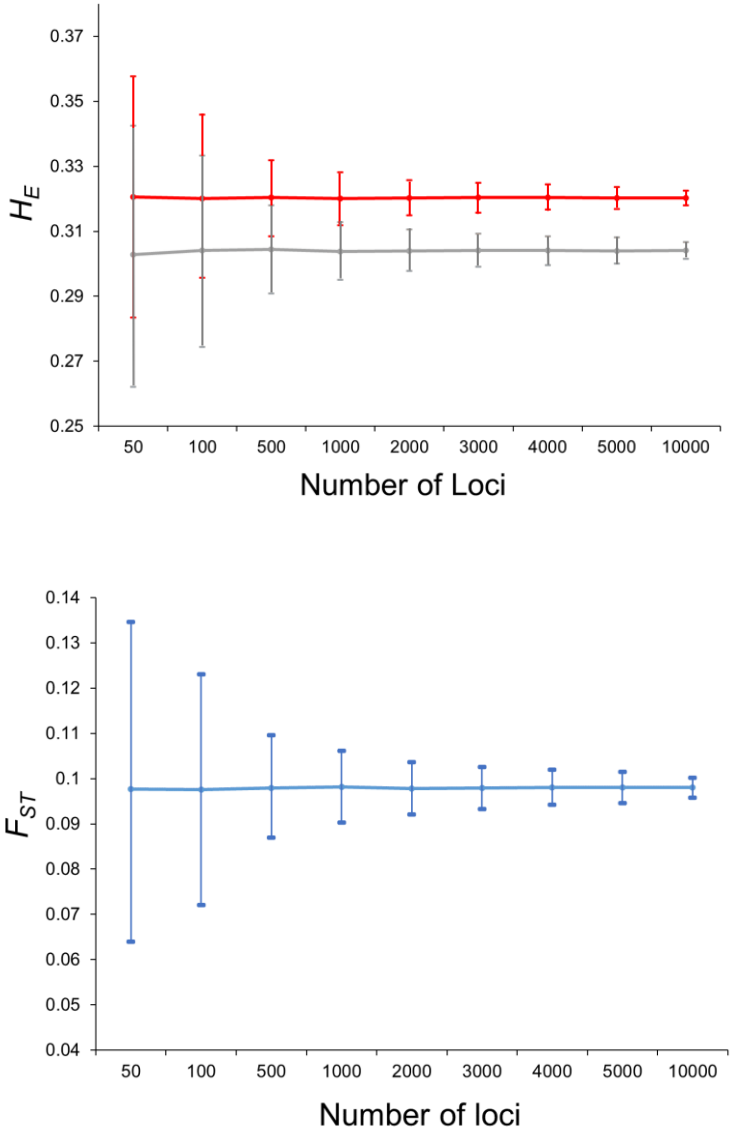

Figure 2.5. Rarefaction analysis results for expected heterozygosity and genetic differentiation as F_{ST} using 3RAD loci and two ponds.

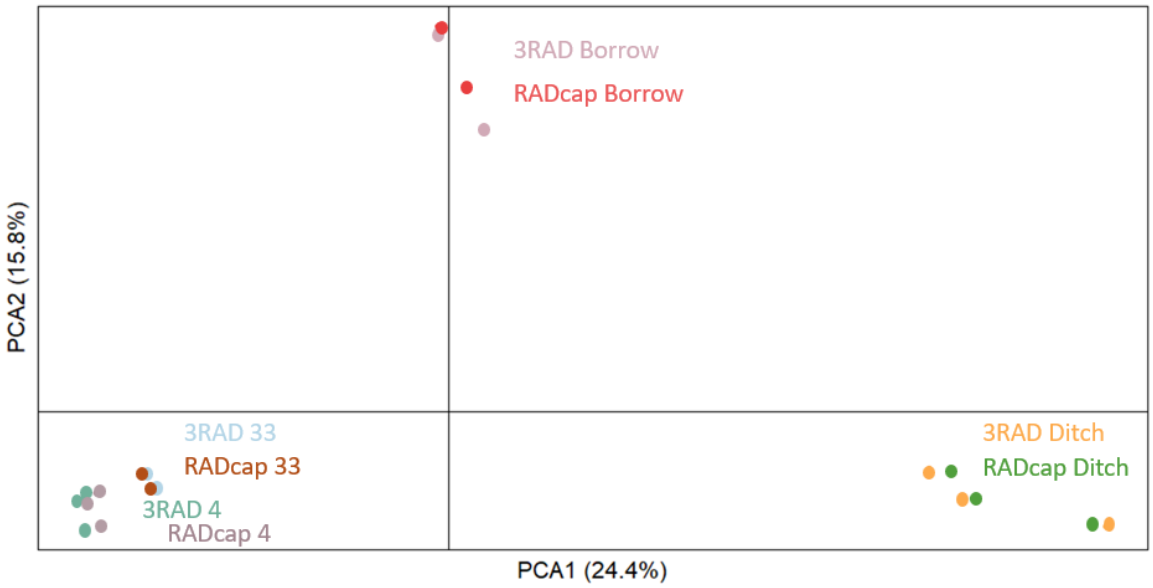

Figure 2.6. Principal components analysis of individuals that were sequenced using both 3RAD and RADcap methods. Points are colored by pond of origin and by sequencing method.

CHAPTER 3

A CONSERVATION GENOMIC INVESTIGATION TO GUIDE MANAGEMENT OF THE ENDANGERED RETICULATED FLATWOODS SALAMANDER (AMBYSTOMA BISHOPI)

Introduction

At least one third of the world's amphibian species are threatened with extinction (McCallum, 2007; O'Donnell et al., 2017), as a result of climate change, disease, habitat loss, and fragmentation (Grant et al., 2016; McCallum, 2007). These stressors alter the demographics and evolutionary trajectories of populations by reducing population sizes, connectivity, and stability (Fahrig and Merriam 1994; Stuart et al. 2004; Zamudio 2007; Graeter et al. 2008). Reduction in population size, combined with increased isolation, decreases resiliency by increasing a population's risk of inbreeding depression, loss of adaptive potential, and extirpation (Ellstrand and Elam 1993; Fischer and Matthies 1998; Jacquemyn et al. 2002; Purrenhage et al. 2009; Greenwald 2010). Loss of genetic diversity, isolation, and extirpation of populations in turn disrupt the larger metapopulation dynamics through reduced redundancy of populations affecting gene flow and source-sink dynamics (Young et al. 1996; Whitlock and Barton 1997; Gonzalez et al. 1998; Pearman and Garner 2006; Spear et al. 2005). These risks associated with habitat loss and fragmentation should be greater for species with specialized habitat requirements and limited dispersal abilities, due to their inability to traverse longer distances through unsuitable habitat in order to maintain gene flow within the metapopulation. As populations become less resilient and the number of populations decreases, the species as a whole experiences declines in genetic and life history diversity (i.e., diversity in age classes, timing of migration, and timing of spawning; Schindler et al., 2010). In order to effectively conserve species facing these threats, management plans must ensure that a) individual populations are resilient to local extirpation through adequate genetic diversity and population size, b)

populations are redundant in that there are enough resilient populations to prevent species extinction, and c) these populations comprehensively represent the genetic and demographic diversity of the species in order to conserve the species' evolutionary legacy. These 3R's – resiliency, redundancy, and representation –have become the cornerstone of endangered species management (Wolf et al., 2015).

The reticulated flatwoods salamander (*Ambystoma bishopi*) is a federally endangered species that faces risks of loss of genetic diversity, inbreeding depression, extirpation, and associated impacts on metapopulation dynamics. *A. bishopi* is endemic to the longleaf-savanna ecosystem of the southeastern United States (Palis 1997). It lays eggs terrestrially in the beds of ephemeral ponds, which hatch after pond inundation, and has a fossorial adult life stage (Gorman et al. 2014; Palis 1995; Palis 1997; Chandler et al. 2016). This complex life history makes it challenging to obtain reliable estimates of population sizes and dispersal through traditional field techniques such as mark recapture of adults. Previous ecological and genetic studies have indicated major range declines and shown that persistence and size of remaining populations is linked to breeding-pond connectivity and habitat suitability, which have been reduced by land-use change, fire suppression, and altered hydrology due to climate change (Pauly et al. 2007; Semlitsch et al. 2017; Van Lear et al. 2005; Brooks et al. 2019a; Wendt et al. 2021). The currently known, extant breeding ponds of *A. bishopi* exist as isolated metapopulations in a few geographic regions within its historic range (Semlitsch et al., 2017). Of these, Eglin Air Force Base (hereafter "Eglin") and Escribano Wildlife Management Area (hereafter "Escribano") currently house the majority of known remaining *A. bishopi* populations, acting as the largest remaining strongholds of the species (USFWS, 2021), but only the Eglin region has received substantial ecological and genetic study in the past. For example, a previous population genetic

study of *A. bishopi* on Eglin by Wendt et al. (2021) in 2013-2016 found that two disparate flatwoods habitat patches on Eglin (Oglesby and Eastbay), functioned as metapopulations, with individual ponds acting as genetically distinguishable subpopulations and gene flow between ponds within flatwoods decreasing linearly and sharply with increased geographic distance. Wendt et al. (2021) also found that Eglin ponds typically had a small number of effective breeders per pond (N_b typically <40 individuals per year) and posited that ponds depend on between-pond habitat connectivity to maintain population persistence and counteract the effects of genetic drift. Continued annual monitoring of population status through estimates of N_b is especially important to the conservation of species where inbreeding depression and population decline are of particular concern, because early detection of N_b decline can help prevent further loss of genetic diversity and population extirpation (Schindler et al., 2010; Schwartz, Luikart, & Waples, 2007). Although our understanding of the species is growing, gaps in *A. bishopi* information remain to be filled, including a) a more range-wide perspective on genetic relationships of major population groups beyond Eglin, b) demographic and genetic diversity trends over time, c) information on genetic diversity, gene flow, and N_b for populations outside of Eglin, d) key environmental drivers of population size and genetic diversity, and e) any knowledge on the genetic architecture of local adaptation. Collecting and analyzing genomic data from other known populations of *A. bishopi* over multiple breeding seasons can provide a more thorough understanding of the species' status and as well as aid our understanding of what is influencing genetics and demographics over time and space.

To combat habitat loss and fragmentation and their effects, various restoration and conservation management tactics are employed at Eglin and Escribano. These activities include (a) manual removal of shrubs and trees in fire-suppressed areas and controlled burning to restore

habitat and improve population connectivity, (b) population monitoring through larval counts, and (c) headstarting programs to increase larval survival. From a 3R perspective, resiliency is being addressed through monitoring populations, restoring breeding habitat, and the headstarting of larvae. In headstarting programs at Eglin and Escribano, eggs are collected from dry pond beds, reared in cattle watering tanks until the late larval/early metamorph stage, and then typically released back into source ponds, thus potentially increasing larval survival, reducing reproductive variance, and increasing the stability of population size. Redundancy, on the other hand, could be improved through the translocation and introduction of headstarted larvae into unoccupied ponds, a tactic under consideration by managers and in limited practice as part o fan adaptive management study at Eglin. However, uninformed headstarting, translocation, and reintroduction activities risk increasing inbreeding levels, genetic swamping of populations with highly related individuals, artificial selection, and outbreeding depression (Frankham et al., 2011; Huff et al., 2011; Wang and Ryman, 2001; Araki et al., 2007) all of which could inadvertently undermine the resiliency of local populations. In order to avoid these negative effects, it is essential to determine the extent to which headstarting samples represent the genetic and demographic characteristics of their source ponds. Furthermore, it is unclear the degree that populations are locally adapted to pond-specific conditions. If translocations cause the loss of some of these adaptations, the result would be a decrease in representation.

Conservation-genomic information could greatly improve our understanding of the status of *A. bishopi* populations and how aspects of the environment influence this status. Data from conservation-genomic analysis should allow managers to enhance the effectiveness of programs aimed at improving the 3Rs for this species. The purpose of this study is to fill information gaps in these areas using a broad-scale conservation genomics assessment utilizing targeted single

nucleotide polymorphism (SNP) genomic markers and high-throughput sequencing of individuals from the entire known range of the species, including an in-depth assessment of populations on Eglin and Escribano. Against that background, this study asks three primary questions.

1) At what spatial scales do we see population structure and gene flow? I will determine the appropriate population units for monitoring and recovery (e.g., resiliency, redundancy, and representation units) and assess genetic relationships among those units. I expect that population structure is hierarchical, with ponds acting as distinguishable population units within larger metapopulations. I also predict that there is continuous and relatively gradual genetic isolation by distance between ponds within metapopulations but discrete and relatively strong genetic differentiation among metapopulations. I based these predictions on the hierarchical population structure and IBD relationships observed in two *A. bishopi* metapopulations on Eglin reported by Wendt et al. (2021) using microsatellite markers, as well as the general unsuitability of habitat conditions between major population centers.

2) What are the effective sizes and levels of genetic diversity of local breeding populations, and how do these measures vary over space and time in relation to environmental conditions? Through answering this second question, I will evaluate population resiliency and determine environmental conditions that support larger, more genetically diverse populations of *A. bishopi*. Based on previous work on Eglin, I expect Eglin populations to be relatively small (small N_b), but that populations in other regions may tend to be larger or smaller, depending on the relative quality and size of habitat patches. I further expect that in addition to breeding-pond size, pond hydrology and connectivity to other occupied ponds will be important determinants of population size and genetic diversity.

3) How are headstarting programs influencing pond demographics and genetic diversity? To answer this last question, I first compare the demographic and genetic characteristics of ponds used in the headstarting program ("headstart" ponds) to those that have not been used in the headstarting program ("wild" ponds), to see if there are significant demographic or genetic differences between these two groups. I then examine pond demographic and genetic characteristics over time to see whether and how headstarting programs seem to be impacting these metrics across breeding seasons. I expected results to support one of three equally plausible hypotheses: 1) headstarting has no apparent effect on demographic and genetic characteristics (metrics are the same between treatment groups and there are no distinguishable trends over time), 2) headstarting has a positive effect by increasing individual survival rate (headstart ponds exhibit larger N_b, lower reproductive variance, and greater genetic diversity than wild ponds, and these metrics are increasing over time), or 3) headstarting actually has a negative effect, potentially by bottlenecking diversity (headstart ponds exhibit smaller N_b, greater reproductive variance, and less genetic diversity than wild ponds, and these metrics are decreasing over time). Lastly, within headstart ponds, I quantify and compare the demographic and genetic characteristics of individuals captured as eggs and then hatched and reared in cattle tanks to those that were "missed" during initial sampling events (either eggs were present but not observed or were laid subsequent to sampling) but subsequently wild-hatched and reared in the origin ponds. This last objective is used to assess the extent to which headstarting could underrepresent the reproductive output, family diversity, and genetic diversity of a pond's cohort by subsampling a fraction of its composition. Within this last objective, I expected results to support one of two equally plausible hypotheses: 1) headstart samples are a nearly comprehensive or at least unbiased subsample of their origin ponds (demographic and genetic

characteristics are the same between headstart samples and the total pond), which would suggest headstarting programs show little risk of shifting pond diversity, or 2) headstart samples represent a substantially reduced and/or biased fraction of the pond's total genetic diversity and demographic characteristics (e.g., number and evenness of breeding families), which would suggest a risk for headstarting programs to constrain pond diversity.

Methods

The data analyzed in this chapter are a result of the laboratory and bioinformatic methods and analyses described in Chapter 2. The dataset involves 4,783 neutral SNP loci sequenced across 1,588 individuals. These individuals represent breeding ponds on Mayhaw, Garcon, Escribano, and Eglin, sampled across five breeding seasons (2016-2021; Figure 2.1). Because *A. bishopi* typically lay eggs in the fall, but larvae do not metamorphose and emerge from ponds until the following spring, each breeding season spans two calendar years, but for simplicity I hereafter refer to each breeding season by the first year involved (i.e., the fall 2019-spring 2020 breeding season is referred to as "2019"). At Mayhaw and Garcon, individuals were sampled from only one breeding pond per property. In contrast, 15 breeding ponds were sampled on Escribano (Borrow, Banana, Ditch, Ghost, Gum, Honey, Restoration, Stanley, Tadpole, Torpedo, EP1, EP5, EP15, EP46, and EP47) and 22 breeding ponds were sampled from the greater Eglin-Hurlburt administrative unit. The latter included 10 ponds on the Eastbay flatwoods region of Eglin (P15, P16, P19, P30, P32, P33, P34, P112, P215, P234), eight ponds on the Oglesby flatwoods region of Eglin (P4, P5, P49, P51, P52, P53, P212, P213), and four ponds on the Hurlburt flatwoods region, adjacent to Eglin and roughly equidistant from Eastbay and Oglesby (H4, H5, H6, H8).

I performed an Analysis of Molecular Variance (AMOVA) in the *poppr* package (Kamvar et al., 2014) in R (RStudio Team, 2020) using 1,000 permutations to determine 1) the magnitude of temporal (between-cohort) genetic variation relative to spatial (among pond) variation, and 2) how much spatial population structure emerged at each of three hierarchical levels: between states (Georgia vs. Florida), among regions within Florida (Garcon vs. Escribano vs. Eastbay vs. Oglesby vs. Hurlburt), and among ponds within regions. Rarefaction analyses were conducted to determine the minimum sample size needed per population to accurately and precisely estimate expected heterozygosity (H_E) and genetic differentiation (F_{ST}) in downstream analyses. For this analysis, four cohorts with relatively large sample sizes were subsampled, including two with relatively small F_{ST} (Torpedo 2019 and Honey 2019) and two with relatively large F_{ST} (Ditch 2019 and P212 2019). Various levels of subsampling (2, 3, 5, 8, 10, 15, 20, 25, 30, 35, and 40 individuals) were evaluated, each replicated 100 times, and bias and precision evaluated by comparing means and 95% confidence intervals to estimates from the "full" dataset, assuming the latter to represent the true values.

As an index of genetic differentiation among populations, I estimated F_{ST} (Weir and Cockerham, 1984) between each pair of ponds using the R package *hierfstat* (Goudet, 2005). From these estimates, I tabulated the average F_{ST} between ponds within regions, between regions (Eastbay vs. Oglesby vs. Hurlburt vs. Escribano), and between properties (Eglin+Hurlburt vs. Escribano vs. Garcon vs. Mayhaw). These F_{ST} estimates were also used to assess IBD (isolation-by-distance; see below). To visualize population relationships based on F_{ST}, I conducted principal coordinates analyses (PCoAs) in the *ape* package (Paradis and Schliep, 2019) for R, using F_{ST} matrices as the response variables. I conducted PCoA at four different spatial extents: a) all ponds included, b) ponds in Florida only, c) ponds on Escribano and Eglin only, and d)

ponds on Escribano only. These groups were analyzed in order to assess population structure between and within properties. Only PCoA axes that showed clear spatial variation (typically only axes one and two) were interpreted.

To get an alternative view of population structure that did not rely on *a priori* population groupings, I used Bayesian admixture analyses in software STRUCTURE version 2.3 (Pritchard et al., 2010) to determine 1) the most likely number (*K*) of hypothetical ancestral populations that gave rise to the data, and 2) how much of each individual's ancestry was drawn from each of these *K* populations. When analyzing highly divergent populations, separate analysis may improve the effectiveness of STRUCTURE in identifying population sub-structure which cannot be detected in a one-step analysis (Evanno et al., 2005; Pritchard et al., 2010). For this reason, a multi-step STRUCTURE approach was performed. An initial run of STRUCTURE was used to identify the top level of population separation followed by separate STRUCTURE analyses within each of these main clusters. This multi-step approach resulted in STRUCTURE runs for a) all ponds from all properties (n = 678), b) ponds on Escribano plus P234 from Eglin (n = 302), and c) ponds from Eglin excluding P234 (n = 368). Due to STRUCTURE's tendency to over-split well-sampled populations and under-split poorly-sampled populations when sample sizes are uneven (Puechmaille 2016), no more than 30 randomly-selected individuals per pond were used in this analysis. For every STRUCTURE analysis, I ran ten replicate models for each assumed *K* value using a burn in period of 100,000 and then 100,000 sampled Markov chain Monte Carlo (MCMC) iterations. I compared likelihoods of *K* values from 1-10 for the range-wide analysis and *K* values from 1-20 for the Escribano and Eglin analyses. Log-likelihood values, delta-*K* values (Evanno et al., 2005), and visual interpretation of bar plots were used to assess and select *K* values that best represented primary population structure for each group

(Faubet et al., 2007). Final *K* values from STRUCTURE were compared to the PCoA patterns and F_{ST} values to compare the delineation of genetic population boundaries between the methods.

Isolation-by-distance (IBD) is a pattern of genetic differentiation whereby individuals and populations that are geographically closer together are more similar genetically than individuals and populations that are geographically farther apart (Hedrick, 2005). Characterizing IBD can aid in the visualization of continuous population structure that is challenging to detect by discrete methods such as AMOVA and STRUCTURE, and further allows one to estimate the spatial scale over which gene flow becomes limited and population structure begins to emerge (Hutchison and Templeton, 1999). Pairwise F_{ST} estimates between ponds were regressed on Euclidean spatial distances between ponds using multiple regression on distance matrices (MRDM) models in package *ecodist* (Goslee et al., 2020) for R. Each MRDM estimated the slope and intercept of the IBD relationship and tested whether the slope was equal to zero based on a permutation test with 10^4 random permutations. I ran regressions at multiple spatial scales, to characterize IBD relationships between states, between regions, and between ponds within regions. For IBD models including multiple regions, I indexed spatial distance using the center-to-center distances between ponds. In contrast, for estimating IBD within individual regions, I indexed spatial distance using the nearest-edge distances between ponds, as the latter type of information was available for Eglin and Escribano and presumably provides a more biologically meaningful measurement of movement cost.

Using ponds as the primary grain of population structure, I estimated N_b, family structure, and several genetic diversity statistics for each pond, to 1) characterize broad regional patterns in genetic diversity, 2) develop point estimates and trends of diversity and N_b, to evaluate the

resiliency of populations, 3) characterize diversity-environment relationships, and 4) assess potential demographic and genetic effects of headstarting. The number of effective breeders (N_b) was estimated for each cohort sampled from each pond using two complementary approaches—linkage disequilibrium (LD) implemented in NeEstimator v2 (Do et al., 2014) and sibship-frequency (SF) using the pedigree approach implemented in Colony 2.0 (Jones and Wang 2010). Random mating, non-overlapping generations, and sampling of a single cohort are assumptions of NeEstimator. *Ambystoma bishopi* is iteroparous; therefore, the assumption of non-overlapping generations is violated. However, all populations should be affected by this violation thus any resulting bias should be similar across all populations (Robinson and Moyer 2013). Nonetheless, I applied the bias-correction formula in Waples et al. (2014) to correct my raw estimates of $N_{b\text{-}LD}$ for an age-structured species like *A. bishopi*. To estimate $N_{b\text{-}SF}$, SNPs were filtered using VCFtools to include only those with a minor allele frequency greater than 0.1 and missing in ≤ 3% of individuals in order to improve resolving power. For analytical tractability, 1000 loci were randomly chosen from this dataset to run in Colony software. Settings in Colony were set to allow polygamous mating for males and females and allow inbreeding but not clones. I used the full likelihood setting combined with the pairwise-likelihood score (FLPS) analysis method with high precision and a medium run length. Sibship scaling was allowed to avoid over splitting large families. No sibship prior and the "update allele frequency" option were applied. A genotyping error rate of up to 0.05 per marker was allowed as allowing for errors prevents erroneous splitting of siblings and generally results in more accurate N_b estimates (Jones and Wang, 2010; Ackerman et al., 2017). As suggested by Waples (2016) and to reduce variance, for each cohort, I calculated the unweighted harmonic mean of N_b ($N_{b\text{-}mean}$) across estimates from NeEstimator *($N_{b\text{-}LD}$)* and Colony *($N_{b\text{-}SF}$)*. To compare SNP-based estimates of N_b with

microsatellite-based estimates, and to include a longer time series of data, $N_{b\text{-}mean}$ was also calculated from Wendt et al.'s (2021) microsatellite data from 2013 and 2015 at Eglin, by running NeEstimator and Colony using the settings described above.

Observed heterozygosity (H_O), H_E, and rarefied allelic richness (A_R where R is the minimum sample size among populations being compared) were estimated for each pond (pooled years) and each cohort within each pond using the *hierfstat* package in R. I used Hohenlohe et al.'s, (2010) method of estimating H_E (i.e., π) instead of the typical method based on Hardy-Weinberg frequencies, as the former is more accurately estimated at small sample sizes. Colony pedigrees were used to determine the number of full-sibling families for each cohort, an index of reproductive success. Full-sibling families determined through pedigree analysis also were used to calculate family evenness (*FE*), which I assumed was inversely related to reproductive variance in family size (Whiteley et al. 2013; Whiteley et al., 2015). *FE* was calculated as $FE = H'/H'_{Max}$, where $H' = \sum_1^S p_i \ln(p_i)$ and $H'_{Max} = \ln(S)$ (Mulder et al. 2004). Here, S was the number of full-sibling families and p_i was the proportion of individuals belonging to the i-th family. Number of individuals per pond, number of polymorphic loci, H_O, H_E, and A_R, were averaged across ponds for each property (Garcon, Escribano, Mayhaw) and flatwoods region within Eglin (Eastbay, Oglesby, and Hurlburt) to compare genetic diversity between properties and flatwoods regions.

In order to identify and assess pond-level drivers of N_b and H_E, I used Pearson correlations to test for relationships with total pond area (ha; Brooks et al. 2019), suitable breeding habitat area (ha; Brooks et al. 2019), pond recession rate (unitless; Chandler et al. 2017), pond hydroperiod (days per year; Chandler et al. 2017), and mean distance to other occupied ponds (km). These analyses were restricted to Eglin ponds, the only region where these

environmental data were available. Correlations were performed using Excel and two-tailed tests to obtain *p* values.

In order to assess how headstart programs may have influenced the genetic and demographic characteristics of ponds, N_b, H_O, H_E, and *FE* were compared between headstart and wild ponds using Wilcoxon signed-rank tests. Comparisons were made using estimates from the 2019 breeding season only, to control variability due to different breeding seasons and because most ponds had sufficient samples sizes in 2019. Estimates of H_E, *FE*, and number of full sibling families were plotted over a series of three or more breeding season where data was available, to assess the potential demographic and genetic impact of headstarting programs over time.

For Escribano ponds and years with sufficient sample size (≥7 individuals per comparison group) and where headstarting was conducted, I assessed the degree to which individuals collected for headstarting, which were subsampled from the total number of eggs laid in each pond and year, represented the demographic and genetic characteristics of the total cohort. This involved comparison of three different datasets for each applicable pond and year: 1) data from individuals collected as eggs during an initial round of sampling then subsequently hatched and reared in cattle tanks ("headstart" samples), 2) data from individuals that were missed during this initial round of sampling, either because eggs avoided detection or were laid subsequent to sampling, and as a result were hatched in the wild and sampled as larvae during a second round of sampling ("wild" individuals), and 3) the "total" dataset comprising the headstart and wild individuals pooled together. The headstart subsample was assumed to be an unbiased representation of what would be represented in a typical headstarting initiative, whereas the total sample was assumed to be an unbiased representation of the overall reproductive output at that pond and year. I compared H_O, H_E, the number of private alleles, number of full-sibling

families, *FE*, and $N_{b\text{-}mean}$ of headstart subsamples to that of the total cohort. The number of private alleles was obtained using the "private_alleles" function in the R package *poppr*. Estimates for captively reared cohorts were then compared to those of naturally reared samples and the entire pond, applying a Wilcoxon signed-rank test when appropriate.

Results

Spatial scaling of population structure and gene flow

I used an AMOVA to evaluate the spatial scaling of population structure and assess whether temporal variation among cohorts was substantial relative to spatial variation. The AMOVA indicated that the genetic variation among cohorts was significantly greater than zero (p <0.001; Table 3.1) but minor (1.4%) relative to spatial variation among ponds (48.7% total across the three spatial strata), so cohorts were pooled by pond for subsequent population structure analyses. The largest spatial difference (35.3% of total variance) was between states (i.e. Mayhaw vs. all other ponds) with progressively less variation attributable to regions within Florida (7.8%) and among ponds in the same region (5.6%). Thus, inferred population structure was strongly hierarchical, but significant even at the spatial grain of individual ponds. These results indicate that a) genetic variation was more attributable to separation in space rather than time, b) population structure was hierarchical, with differentiation increasing with spatial scale, and c) Mayhaw was by far the most genetically differentiated location, but d) even at the grain of individual ponds, population structure was meaningful and substantial.

Unlike AMOVA, STRUCTURE uses no *a priori* information about spatial sampling locations, but uses a Bayesian clustering method to determine the number of genetic clusters that gave rise to the dataset. STRUCTURE results agreed with those of AMOVA in that population structure was indicated to be hierarchical and that population structure often was detectable at the

individual pond level. For the full dataset, a model of 10 populations (K=10) had the highest average log-likelihood (appendix), but created non-meaningful subdivisions (i.e., not corresponding to geography or time) within and between ponds (not shown). Alternatively, the ΔK method proposed by Evanno et al. (2005) would indicate a K of 2 as the uppermost level of structure (appendix), but such a model would leave clear, geographically meaningful structure undescribed, which was increasingly captured as K increased from 2 to 6 (Figure 3.3). I therefore interpreted the K=6 model as the best representation of top-level population structure for *A. bishopi*. Based on this model, the primary genetic distinction was between Escribano and Eglin, along with a secondary distinction between the Oglesby and Eastbay regions on Eglin and some distinctiveness of other individual ponds including Ditch, Stanley, Borrow, and P212 (Figure 3.3). Unexpectedly, Mayhaw and Garcon did not constitute unique genetic lineages, but rather appeared as admixtures of the Eglin and Escribano lineages. Furthermore, Hurlburt showed more genetic relatedness to Escribano than expected based on the spatial juxtaposition of these areas.

Given that the greatest apparent distinction was between Escribano and Eglin (including Hurlburt) (i.e. K=2), I ran separate STRUCTURE sub-analyses within each of these two subgroups. Mayhaw and Garcon were excluded, whereas Eglin pond P234 was included with Escribano because it was recognized by STRUCTURE (and other analyses) to be more genetically similar to Escribano. The Escribano sub-analysis indicated a model of 20 populations (K=20) to have the highest average log-likelihood and a model of K=2 to have the highest ΔK (appendix). However, K=20 clearly over-split ponds and created non-geographically-meaningful divisions (not shown), whereas K=2 failed to capture apparently meaningful population structure that was explained by successively higher-K models (Figure 3.4). I therefore selected the K=5 model as the best representation of population structure within this dataset. Based on this model,

four ponds or pond-groups (Borrow plus P234, Ditch plus Stanley, EP15, and EP46) were differentiated from the other ponds on Escribano (Figure 3.4).

STRUCTURE analysis of Eglin and Hurlburt ponds indicated K=19 as the uppermost level of structure based on log-likelihood and K=2 based on ΔK (appendix), but based on visual interpretation of STRUCTURE bar plots, K=4 seems to best characterize geographically and genetically meaningful divisions between ponds on Eglin (Figure 3.5). The K=4 model essentially separated the three main flatwoods regions, Eastbay, Oglesby, and Hurlburt, and within Oglesby, pond P212 was further distinguished from other ponds. Unexpectedly, pond P19 from Eastbay grouped with Hurlburt ponds.

I estimated pairwise F_{ST} between ponds to further quantify genetic differentiation and, by extension, levels of connectivity between ponds. On average, the greatest genetic distance was between Mayhaw and ponds from Florida (average F_{ST}=0.34; Table 3.2) followed by the genetic distances between properties on Florida (average F_{ST} between ponds: Garcon vs Escribano=0.19, Garcon vs Eglin=0.18, Escribano vs Eglin=0.14). The next greatest genetic distance was between Eastbay and Oglesby regions on Eglin (average F_{ST}=0.10). The least amount of genetic differentiation was within flatwoods regions: Escribano (average F_{ST}=0.08), Oglesby (average F_{ST}=0.07), Eastbay (average F_{ST}=0.07), and Hurlburt (average F_{ST}=0.04). Interestingly, average genetic differentiation between Hurlburt and Escribano (average F_{ST}=0.08) and between Hurlburt and Eglin (average F_{ST}=0.08) was no greater than differentiation within Eastbay or Oglesby. Rarefaction analysis was conducted to determine the number of individuals required to accurately estimate F_{ST} between ponds. Rarefaction results indicated a downward bias and reduced precision (wider confidence intervals) in F_{ST} estimates when using <8 individuals

compared to ≥8 individuals (Figure 3.2a). Therefore, only ponds with ≥8 individuals were included in subsequent F_{ST} analyses.

I used pairwise F_{ST} values in which n ≥8 individuals for both ponds to conduct a PCoA and visualize continuous population structure among ponds. PCoA results largely agreed with the findings of AMOVA and STRUCTURE in that genetic distinctions between populations corresponded to geographic boundaries of properties and regions, and yet population substructure was seen within flatwoods regions as well. The range-wide PCoA including all sites (32 ponds) showed clear separation between Mayhaw and all Florida ponds, as well as among Florida properties (Garcon, Escribano, and Eglin), but did not show an expected distinction between Oglesby and Eastbay on Eglin (Figure 3.6a). The clear separation between Mayhaw and Florida ponds supported AMOVA findings that the greatest genetic distinction was between states. Separation of Florida properties also supported findings of STRUCTURE and AMOVA in that population structure seemed to be hierarchical. Separate PCoAs were conducted using subsets of ponds to visualize population structure at finer scales. The PcoA using ponds from Escribano and Eglin only (30 ponds) showed separation between flatwoods regions, with Hurlburt appearing genetically intermediate to Escribano, Eastbay, and Oglesby ponds (Figure 3.6c). This clustering pattern supported the unexpected amount of shared genetic lineage between Escribano and Hurlburt in the range-wide STRUCTURE results. The PcoA including only Escribano ponds plus Eglin pond P234 (14 ponds) showed clear separation of Borrow plus P234 from other ponds on Escribano (Figure 3.6d). The Escribano PCoA also showed Ditch plus Stanley separating from the other ponds on the property. This separation on the pond level further supported that population structure was hierarchical with ponds acting as the smallest distinguishable population units.

To quantify the spatial scaling of gene flow within and among metapopulations, I analyzed IBD relationships between pairwise F_{ST} and geographic distance at the grain of individual ponds, but at varying spatial extents. These analyses showed a significant, positive IBD relationship at all extents considered, but the slope and intercept of the relationship varied strongly with extent. The range-wide analysis of IBD including all sites with ≥8 individuals (32 ponds) exhibited a significant, positive IBD relationship (F_{ST} =0.001(km) + 0.0897, R^2=0.61, p=0.0001) and primarily showed substantial differentiation between Georgia (Mayhaw) and Florida ponds (Figure 3.7). When Mayhaw was excluded and only ponds in Florida were analyzed (31 ponds), there was a noticeable difference in IBD trends at spatial scales less than 5 km vs greater than 5 km. Among ponds <5 km apart, there was a strong, significant, positive IBD relationship between ponds, with F_{ST} increasing by about 0.025 for every 1 km of distance (F_{ST} =0.0252(km) + 0.0398; R^2=0.50; p=0.0001). Among ponds greater than 5 km apart, the IBD relationship was still significant (p=0.0242), but the relationship was much weaker, with F_{ST} increasing by 0.001 for every 1 km (F_{ST} =0.001(km) + 0.0991; R^2=0.04). Looking only at Eglin and Escribano ponds within 5 km of each other, IBD relationships were positive and significant on both Eglin (F_{ST}=0.036(km) + 0.0442; R^2=0.53; p=0.0008) and Escribano (F_{ST}=0.024(km) + 0.0423 ; R^2=0.51 ; p=0.0001) and exhibited similar slopes and intercepts (Figure 3.8). For Eglin, F_{ST} increased by 0.036 for every 1 km, whereas for Escribano, F_{ST} increased by 0.024 for every 1 km. Very small F_{ST} values (<0.01) occurred only between ponds separated by less than ~500 m, whereas moderately small F_{ST} values (<0.05) occurred only between ponds separated by less than ~1.2 km (Figure 3.8).

Size and Genetic Diversity of populations

As an index of population size, recruitment, and resiliency, I estimated the number of effective breeders (N_b) that produced each cohort in each pond by taking the harmonic mean ($N_{b\text{-}mean}$) of Colony (Jones and Wang 2010; $N_{b\text{-}SF}$) and adjusted NeEstimator results (Do et al., 2014; Waples et al., 2014; $N_{b\text{-}LD}$). Estimates of N_b from microsatellite data (Wendt et al., 2021) appeared comparable to SNP-based estimates (Table 3.3), so I co-analyzed these results. There was a tendency for $N_{b\text{-}SF}$ to exceed $N_{b\text{-}LD}$ (71% of cases), with $N_{b\text{-}SF}$ estimating an average of 8 more individuals per cohort than $N_{b\text{-}LD}$. However, rank orders of the two methods were strongly positively correlated (Spearman's *rho*=0.72, $p<.0005$), indicating that they similarly discriminated high- from low-N_b cohorts. Only cohorts with ≥10 sampled individuals were used to estimate $N_{b\text{-}mean}$, because smaller samples resulted in inestimable pedigrees (Colony) or infinite estimates (NeEstimator). Of the 45 pond-year cohorts in which $N_{b\text{-}mean}$ was estimable, estimates ranged from 4 to 104, was <50 individuals in 80% of instances, had a mean of 26 and a median of 21 (Table 3.3). Average $N_{b\text{-}mean}$ was similar on Eglin (23 individuals) and Escribano (28 individuals), with Escribano showing a wider range of estimates (4-104) compared to Eglin (4-68; Table 3.3). Among ponds for which multiple cohorts were sampled, temporal trends in $N_{b\text{-}mean}$ appeared temporally correlated across ponds both between and within regions (Figure 3.9). Ponds Honey, Torpedo, Ditch, and Borrow experienced a decrease in $N_{b\text{-}mean}$ between 2018 and 2019, and all but Ditch decreased again between 2019 and 2020 (Figure 3.9). In contrast, Ghost's $N_{b\text{-}mean}$ increased between 2018 and 2019 and between 2019 and 2020 (Figure 3.9). Similarly, Eglin ponds P15, P53, and P215 all saw a decrease of $N_{b\text{-}mean}$ between 2019 and 2020, while ponds P4, P5, and P212 all saw an increase of $N_{b\text{-}mean}$ between 2013 and 2015 (though P15 decreased over the latter time period).

In order to monitor and compare neutral genetic diversity among ponds and flatwoods regions, I estimated H_O, H_E, and A_R, for each pond and tabulated the number of private alleles for each property. Rarefaction results indicated an upward bias and reduced precision of H_E estimates with <8 individuals compared to ≥8 individuals sampled per population (Figure 3.2b). Therefore, diversity statistics were estimated and presented only for ponds with ≥8 sampled individuals. Across all ponds, average expected and observed heterozygosity were identical (0.31) and exhibited similar ranges (H_E=0.28-0.34, H_O=0.27-0.36; Table 3.4). H_E, H_O, and A_R were averaged for each flatwoods region, and because Garcon and Mayhaw both contain only 4 individuals, ponds from Eglin and Escribano with ≥4 individuals were included in these estimates. Ponds 30 and 49 from Eglin were excluded because they contained translocated individuals from multiple other ponds. Estimates for A_R, H_O, and H_E, were slightly lower on Mayhaw (A_4=1.22, H_O=0.25, H_E=0.22) than in Florida regions. Estimates were about the same between flatwoods regions in Florida with A_4 averaging 1.29 with a range of 0.04, average H_O of 0.37 with a range of 0.07, and average H_E of 0.29 with a range of 0.04 (Table 3.5). At the property level, Escribano contained 8 private alleles and Eglin contained 10 private alleles, whereas neither Mayhaw nor Garcon exhibited private alleles (not shown). However, when comparing regions, there were no Eglin alleles private to Eastbay, Oglesby, or Hurlburt (i.e., all Eglin alleles were shared across at least two of these regions).

To evaluate which environmental features most influence inter-pond variation in both recruitment and neutral genetic diversity, I correlated each of five indices of pond habitat conditions to the $N_{b\text{-}mean}$ and H_E of that pond. Because most of these indices were available only for Eglin, the analysis was restricted to 13 ponds on Eglin that had both habitat data and estimable $N_{b\text{-}mean}$ values. There was a significant, positive correlation between $N_{b\text{-}mean}$ and total

pond area (r=0.58, p=0.014) and a significant, negative correlation between $N_{b\text{-}mean}$ and pond recession rate (r=-0.71, p=0.002; Figure 3.10). The negative correlation between $N_{b\text{-}mean}$ and the average distance (m) to other ponds was moderately strong but not significant (r=-0.44, p=0.078). Weaker, non-significant correlations were found between $N_{b\text{-}mean}$ and area of suitable breeding habitat, and pond hydroperiod (r=0.28 and 0.29, respectively; both p>0.05). There was a positive correlation between H_E and total pond area (r=0.41, p=0.16) and significant, negative correlations between H_E and the average distance (m) to other ponds (r=-0.65, p=0.02) and pond recession rate (r=-0.75, p=0.003; Figure 3.10). Weaker, non-significant correlations were found between H_E and area of suitable breeding habitat and pond hydroperiod (r=0.29 and 0.25 respectively; both p>0.05).

Demographic and genetic characteristics of headstart samples

To assess how headstart programs may have influenced demographic and genetic diversity at the pond level, 2019 estimates $N_{b\text{-}mean}$, H_O, H_E, and FE were compared between ponds that were used in the headstart program (Borrow, Ditch, Stanley, Torpedo, Honey, Gum, Ghost, EP15, P212, P215, P33, P5, P4, P15, P16, P112, P53) and non-headstarted ponds (Restoration, EP5, EP1, EP46, EP47, P52, P32, P213, H4, H5, H6, H8). I found that headstart ponds tended to exhibit larger N_b, but little difference in genetic diversity, compared to non-headstarted ponds. Neither $N_{b\text{-}mean}$ (t=64.5, p=0.22), H_O (t=82, p=0.34), H_E (t=137, p=0.81), nor FE (t=57.5, p=0.28) was significantly different between treatment groups. Headstarted ponds showed greater average $N_{b\text{-}mean}$ (mean of 33 individuals) than non-headstarted ponds (mean of 12 individuals; Table 3.3). Of the 8 ponds possessing estimated $N_{b\text{-}mean}$ >20, one is a wild pond while the other 7 are headstart ponds. All ponds possessing estimated $N_{b\text{-}mean}$ >50 were headstart ponds (Table 3.3). Ponds with estimated $N_{b\text{-}mean}$ <20 were more evenly split, with 6 headstart

ponds and 7 wild ponds. Headstart pond H_O ranged from 0.27 to 0.35 with an average of 0.31, and H_E ranged from 0.28 to 0.34 with an average of 0.31 (Table 3.4). Wild pond H_O ranged from 0.29 to 0.36 with an average of 0.31, and H_E ranged from 0.28 to 0.33 with an average of 0.31 (Table 3.4). Headstart pond *FE* ranged from 0.79 to 1 with an average of 0.89, whereas wild pond *FE* ranged from 0.80 to 1 and averaged 0.92.

In order to assess how headstart programs may have influenced the genetic and demographic characteristics of ponds over time, H_E, number of full-sibling families, and FE were plotted over time for all headstart ponds with multiple years of data. No wild ponds met these criteria. These statistics were able to be estimated over a time series of three or more years for seven headstart ponds (Borrow, Ditch, Ghost, Honey, Torpedo, P4, and P5; Figure 3.11). Results were ambivalent about the positive or negative influence of headstarting. Comparing the first and last years of sampling, H_E increased in five of seven ponds, but number of full-sibling families decreased in five of seven ponds and *FE* decreased in four of seven ponds (Figure 3.11).

To determine if headstart cohorts effectively represented the genetic and demographic characteristics of origin ponds, diversity and demographic statistics were compared between headstart subsamples and the entire pond of origin (i.e., headstart- plus wild-hatched individuals, combined; Table 3.6). Wilcoxon signed-rank tests showed no significant differences in the amount of neutral genetic diversity (quantified by H_O and H_E) captured in headstart subsamples compared to the entire pond of origin (H_O: t=153.5, p=0.76; H_E: t=137, p=0.81). However, headstart subsamples contained a significantly smaller number of full-sib families than the total (t=77, p=0.02), missing on average 37% (9/20) of families (Table 3.6; Figure 3.12). Headstart subsamples also missed an average of 2.1% (190/9,097) of the total alleles (Table 3.6; Figure 3.12). On average, headstart subsamples exhibited an $N_{b\text{-}mean}$ of nine fewer individuals (median

of 5) than the "true" $N_{b\text{-}mean}$ estimated from the total sample (Table 3.6), but this difference was not significant (t= 97, p=0.11). On average, family evenness was slightly higher in headstart subsamples than in the total (average difference in FE=-0.01; Table 3.6), but this difference was not significant (t=168, p=0.43).

Discussion

Spatial scaling of population structure and gene flow

By far, the strongest population structure was between populations from different states, as seen in the AMOVA, F_{ST}, and PCoA results. Based on F_{ST}, Mayhaw was twice as different as any other population, and based on AMOVA, the loss of Mayhaw would mean the loss of 35% of total neutral variation. This finding implies that in order to comprehensively represent the genetic diversity of *A. bishopi* and conserve the species' evolutionary legacy, management goals must include populations on Mayhaw. Failing to represent Georgia *A. bishopi* populations in conservation plans could lead to major loss in *A. bishopi* genetic diversity.

Within properties, ponds clustered into genetically distinct flatwoods regions (Escribano, Eastbay, Oglesby, and Hurlburt), as seen in AMOVA, STRUCTURE, and IBD plots. These results support that flatwoods regions function as separate metapopulations that exchange little to no gene flow in contemporary landscape conditions. Similar population structuring was observed for *A. bishopi* using microsatellite markers from Eastbay and Oglesby ponds in Wendt et al. (2021), and metapopulation structuring has also been reported in other salamander species (Zamudio and Wieczorek, 2007; Sunny et al., 2014; Pisa et al., 2015). This relates to conservation management in that these flatwoods regions should be monitored and managed as independent units when assessing resiliency and redundancy. The exception to this was pond P234 from Eglin's Eastbay region being more genetically similar to nearby ponds on Escribano

than to other, spatially distant ponds on Eastbay. This finding indicates that *A. bishopi* population dynamics do not adhere to these administrative boundaries and highlights the importance of cross-jurisdictional coordination of recovery efforts

At the finest spatial grain, most individual ponds were genetically distinguishable based on F_{ST}, suggesting that ponds function as semi-dependent subpopulations within larger connected regional metapopulations. This semi-dependency of ponds agrees with the genetic findings of previous *Ambystoma* studies to further support generally restricted gene flow (Zamudio and Wieczorek, 2007; Purrenhage et al., 2009; Wang et al., 2009; Wendt et al., 2021). These results support that ponds are genetically distinct units. The noteworthy amount of genetic distinction between ponds (average F_{ST} within flatwoods regions of 0.07) could be a result of philopatry, which is often seen in pond breeding amphibians (Gamble et al., 2007), small effective population sizes increasing the effect of drift, low migration rates due to unsuitable habitat between ponds, adaptive variation between ponds, or a combination of these phenomena. The genetic differences between *A. bishopi* ponds increase with increasing distance as seen in the IBD plots in agreement with the previous findings of Wendt et al. (2021). Increasing genetic difference with increasing distance is most likely an effect of limited dispersal capabilities of the species and specific habitat requirements impacting migration rates between ponds. In light of these findings, ponds should be thought of as the smallest unit for assessing resiliency (i.e. within ponds) and redundancy (i.e. across ponds within regions) and form the foundation for management goals and activities such as restoring and repopulating individual ponds and tracking population metrics on the pond level. Additionally, plans surrounding reintroduction should take into consideration the genetic distinctions between metapopulations. For example, to minimize risks of outbreeding depression and loss of representation, it may be more prudent to

repatriate an unoccupied pond on Eastbay using individuals from other ponds on Eastbay rather than ponds on Oglesby or elsewhere.

Strong differences in the IBD relationship among and within regions suggest that gene flow between *A. bishopi* populations is scale-dependent. This is seen in the strong, significant IBD relationship between ponds within Eglin and Escribano flatwoods regions up to about 5 km that breaks into a still significant but substantially weaker IBD relationship between ponds separated by more than 5 km (i.e., across region boundaries). The weak genetic isolation by distance trend between ponds separated by more than 5 km suggests there has not been enough gene flow between these regions in the recent past to counteract drift. Thus, areas separated by more than 5 km appear to be on independent genetic trajectories (i.e., are functionally different populations), whereas ponds within 5 km are functionally connected through gene flow, similar to "Case IV" in Hutchison and Templeton (1999). This highlights the importance of maintaining or restoring pond connectivity, either through habitat or population restoration, to maintain gene flow within metapopulations, and to the extent possible, between them. This is because with stable conditions over time, gene flow would eventually spread to all degrees of geographic separation and counteract drift ("Case I" in Hutchison and Templeton, 1999). However, with continued dispersal restrictions (caused by pond extirpation and habitat degradation), gene flow would be overpowered by drift causing metapopulations dynamics to collapse.

A. bishopi dispersal seems to be limited within regional metapopulations. At between-pond distances less than 5 km within flatwoods regions on Escribano and Eglin, connectivity decreased consistently with the distance separating ponds, at a rate of ~ 0.02 increase in F_{ST} per km on Escribano and ~ 0.04 increase in F_{ST} per km on Eglin. The only negligible genetic distances occurred at distances less than 500 m, which agrees with previous findings of little to

no migration among ponds on Eglin beyond about 400 m (Wendt et al., 2021). There seemed to be another break in IBD trends at both Eglin and Escribano around 1.2 km, beyond which the minimum observed F_{ST} values increased sharply. This finding corresponds with that of indirect connectivity estimates made by Brooks et al. (2019) based on occupancy models at Eglin. The authors estimated negligible colonization rates between ponds separated by distances greater than 1.5 km. Although my F_{ST} estimates continued to increase with distance beyond 1.5 km, this could be due to the nature of F_{ST} to take many generations (increased with low migration rates) to reach a new equilibrium following fragmentation (Whitlock, 1992), which suggests gene flow between ponds separated by >1.5 km sometime in the past but does not provide strong support for current gene flow at these distances. The similarity of the IBD relationship on both Eglin and Escribano suggests that these dispersal patterns are true for the species as a whole and not unique to Eglin regions. Because *A. bishopi* appears to have a limited dispersal range, the distance from occupied breeding ponds should be considered when determining inactive ponds to receive reintroductions and areas for restoration activities. Further, ponds that are separated from other active ponds by more than 1.5 km may be of particular concern for loss of genetic diversity (see below) and local extinction events due to increased isolation. Restoring the intervening habitat and repopulating inactive ponds could help preserve these separated populations by increasing connectivity. Selecting ponds that are separated from other occupied sites for headstarting programs may allow these sites to persist, and thus aide in preventing local extinction.

Size and Genetic Diversity of populations

The demographic and genetic indices generated by this study contribute to a time series of resiliency indicators that will be valuable for tracking species status, success, and recovery. Tracking estimates of heterozygosity and demography such as N_b and number of families over

time allows us to determine if recruitment, population size, and genetic diversity are increasing, decreasing, or remaining stable (Waples et al., 2013). This information can then be used to evaluate management regimes and recovery progress.

The estimated effective number of breeders per year ($N_{b\text{-}mean}$) varied among ponds and breeding seasons but typically was relatively small for ponds on both Eglin and Escribano, with most ponds possessing $N_{b\text{-}mean}$ between 4 and 68 individuals. A notable exception was the Honey pond on Escribano with an estimated 104 breeding individuals in 2018 and 101 individuals in 2019. Estimates of N_b did not vary substantially between Eglin and Escribano. This consistency in N_b estimates could be a response to the similar management regimes and habitat conditions between properties. Ponds with particularly low N_b on Eglin had smaller total areas and receded quickly. Although correlations were not performed for Escribano ponds, it is notable that Ditch pond had consistently low N_b estimates and a small total area. There appeared to be regional correlation of N_b temporal dynamics, both within and between Escribano and Eglin. Of the eight ponds with estimable N_b for the 2019 and 2020 breeding seasons, six experienced decreases between these years. This decrease corresponds to reports on regional weather patterns, in that pond drying and flooding events provided favorable breeding conditions in the 2019 breeding season and less favorable conditions during the 2020 breeding season due to tropical storms flooding ponds early in the breeding season (preventing egg laying) and less rainfall later in the season causing ponds to dry more quickly (preventing successful metamorphosis and survival). In light of Chandler et al. (2016)'s findings that *A. bishopi* breeding ponds hold water for shorter periods of time in recent years than any other point since 1896, the decrease in N_b corresponding to poor weather conditions further suggests that climate change will negatively impact *A. bishopi* reproductive success through providing unfavorable conditions for metamorphosis. The low N_b

estimates seen in this study (median = 21) are comparable to estimates of N_b across 29 populations of *Ambystoma opacum* (median=42; Whiteley et al., 2014) which has the same breeding strategy as *A. bishopi* (terrestrial deposition as eggs during the late fall months) but is not imperiled. Wang et al. (2011) reported similarly low N_b values across 10 populations of the vulnerable *Ambystoma californiense* (median=29). These values contrast starkly to N_b estimates for 19 populations of the widely-spread, spring-breeding species of least concern, *Ambystoma maculatum* (median=136; Whiteley et al., 2014). This interspecific variation could be caused by a number of differences between these species including habitat requirements, population distributions, and breeding seasons

Despite substantial spatiotemporal variation in N_b, number of full-sibling families, and *FE*, genetic diversity was highly similar across most regions and varied little over time. This suggests that 1) occasional "boom" years, combined with a long lifespan, buffer the genetic effects of more frequent "bust" years, allowing the maintenance of genetic diversity, and b) ponds in the same region are interconnected enough by gene flow over ecological time to maintain genetic diversity over time at the metapopulation scale. The maintenance of metapopulation genetic diversity through regional gene flow indicates the importance of preserving regional dynamics and pond connectivity to conserve genetic diversity. On the other hand, the diversity of Mayhaw was lower than other regions, potentially as a result of spatial isolation preventing genetic rescue from other ponds.

Presumably, all extant *A. bishopi* populations have been fragmented and at least partially bottlenecked by habitat losses throughout the historical range, thus we lack baseline estimates of N_b and H_E levels in the absence of anthropogenic impacts. Further, we do not know the quantitative relationships between population size and extinction risk, probability of inbreeding

depression, or rate of loss of adaptive alleles. Lacking these species-specific resiliency criteria, we can compare results to general rules of thumb, such as the "50:500" rule, which stipulates that a) N_e per generation should be >50 individuals over the short term to avoid inbreeding depression, and b) H_E should decline by <0.1% per generation due to drift (i.e., N_e should be >500) in order to maintain adaptive potential over the long term (Kimura 1955, Wright, 1931; Franklin, 1980). For most ponds in recent years, N_b was below 50 (average N_b =26). Based on age-specific survival, fertility, and reproductive variance, Brooks et al. (unpublished) estimated *A. bishopi*'s N_b:N_e ratio to be 0.81, giving me an average N_e of 32 individuals per generation. Although this is <50 on an individual pond level, the low F_{ST} estimates in this study and the substantial migration rates estimated by Wendt et al. (2021) among nearby ponds suggest that most ponds are not truly isolated over ecological time, such that the collective N_e/N_b of connected ponds would likely exceed 50 on average. Likewise, over the decadal timescales relevant to the maintenance of heterozygosity, the focal gene pool would be an entire connected metapopulation (in this case, a region or property), and over such spatial and temporal extents, the sum of all ponds' multi-generational N_e's may well approach or exceed the 500 threshold. Moreover, there was no apparent declining trend in H_E over time, in ponds where this could be examined, though a longer time series would be needed to formally test for such a decline. Taken together, these findings tell us that *A. bishopi* ponds that are connected to a network of ponds through gene flow have greater resiliency (as measured through N_b and H_E) than more isolated ponds. This implies that more isolated ponds (Mayhaw, Garcon, Ditch, Borrow, P234) are less resilient and more likely to experience the loss of genetic diversity, decreased population sizes, inbreeding depression, and eventually local extinction. Annual pond N_b, number of full-sibling families, and H_E should continue to be monitored as measures of resiliency to ensure population

demographics are periodically recovering from unfavorable breeding seasons and to detect population demographic and genetic declines. The most recent species status assessment for *A. bishopi* evaluated resiliency based on measures of body size and condition and habitat quality at each extant site, in addition to N_e for a limited number of ponds (USFWS 2020). I propose that the addition of pond N_b, number of full-sibling families, H_E, and the number of occupied ponds within a 500 m pond-edge radius to assessment criteria would provide a more accurate evaluation of pond resiliency and functional redundancy.

Observed relationships between N_b, genetic diversity, and pond habitat characteristics could aid managers in prioritizing ponds for protection and restoration. On Eglin, I found that larger breeding ponds that dried slowly and were spatially closer to other occupied ponds tended to exhibit larger N_b and greater H_E than small ponds that dried faster and were more isolated. More isolated ponds are less likely to receive immigrants or colonists for *A. bishopi* and other *Ambystoma* species (Brooks et al., 2019; Gamble et al., 2007; Scott et al., 2013; Wang and Shaffer, 2017). Larger wetlands may be able to support more widespread reproductive success due to increased availability of resources such as vegetation and invertebrate food types (Barber et al., 2004), relative to smaller wetlands. The correlation with pond hydrologic recession rate found in this study may be driven by *A. bishopi's* relatively long development time (11-18 weeks; Palis, 1995). Wendt et al. (2021) found similar relationships between total area, recession rate, mean distance to other ponds and *A. bishopi* genetic diversity estimates using microsatellite markers. However, the only significant correlation of these was between mean distance to other ponds and allelic richness. Significant, positive correlations between N_b or N_e and total pond area have been previously reported for other pond breeding *Ambystoma* species, including *Ambystoma texanum* (Rhoads et al., 2017), *Ambystoma tigrinum* (McCarney-Melstad et al.,

2018), and *Ambystoma californiense* (Wang and Shaffer, 2017; Wang et al., 2011). Rhoads et al. (2017) also reported a significant correlation between N_e and average distance to other sampled ponds of *A. texanum*. Alternatively, Wendt et al. (2021) found a positive, significant correlation between *A. bishopi* pond area of suitable breeding habitat and microsatellite estimates of H_E. In contrast, I did not find expected relationships of N_b or H_E with the area of suitable breeding habitat or pond hydroperiod. The variation between my environmental findings and those of Wendt et al. (2021) is likely a result of differences in the type and number of data markers (9 microsatellite loci vs 4,783 SNP loci). Several studies have reported weak correlation between SNP- and microsatellite-based estimates of genomic diversity and more resolving power of large SNP data sets estimates of genomic diversity with larger SNP data sets (Lemopoulos et al., 2019; Fischer et al., 2017; Camacho-Sanchez et al., 2020). Additional sources of variation could be sample size differences (Pruett and Winker, 2008; Hale et al., 2021) and the use of pooled breeding seasons for H_E estimates in Wendt et al. (2021). Based on these findings, managers seeking to increase population resiliency through increasing N_b and H_E might focus on restoration, such as hardwood removal (Golladay et al., 2021), to increase pond area and decrease recession rate as well as target large, slow drying ponds that are close to other occupied ponds for reintroduction events.

Demographic and genetic characteristics of headstart samples

When investigating if diversity is detectably changing over time in headstart ponds, I expected one of three possible outcomes: 1) no detectable increase or decrease in genetic diversity over time (headstarting has no effect), 2) a downward trajectory in genetic diversity (headstarting is bottlenecking diversity), or 3) an upward trajectory in genetic diversity (headstarting is increasing reproductive success and genetic diversity). Based on the data at hand,

headstarting did not have a consistent positive or negative effect on genetic diversity or pond demographic characteristics, though there is only a limited time series of data so far with which to test these hypotheses. Thus, there is no evidence yet to conclude whether headstarting is having a genetic or demographic effect on *A. bishopi* populations. This conclusion may change, given a longer time series of data and when compared to wild pond diversity and demographics over time.

If headstarting is reducing reproductive variance, increasing juvenile survival, and thereby increasing demographic and genetic diversity, I would expect headstart ponds to maintain greater heterozygosity, larger effective population sizes, and a greater number and evenness of families compared to wild ponds. Unfortunately, too few wild ponds exhibited enough temporal data to allow a comparison of trends, so static comparisons were made between headstart and wild ponds. There was no evidence that headstart ponds significantly increased genetic diversity relative to non-headstart ponds. On the other hand, headstart ponds did exhibit greater N_b and a greater number of families than wild ponds on average. However, these differences should be interpreted with caution, given that ponds were not assigned to these two "treatments" randomly. Rather, this was an unplanned experiment, in which managers had pre-selected which ponds to use for the headstarting program based on other factors, which likely included pond size, accessibility, and the ability to collect large numbers of eggs and larvae for cattle tanks. As such, there may be a built-in bias for headstart ponds to have larger numbers of breeders and greater recruitment on average, even in the absence of the headstarting program.

When assessing the extent to which headstart cohorts represented the demographic and genetic characteristics of their populations of origin, I expected one of two outcomes: 1) headstart samples are a nearly comprehensive or at least unbiased subsample of their origin

ponds (i.e., demographic and genetic characteristics are the same between headstart samples and the total pond), or 2) headstart samples represent a substantially reduced and/or biased fraction of the pond's total genetic diversity and demographic characteristics. To achieve its goals, headstarting should increase the number of surviving families and number of surviving offspring within each family, *without* disproportionately favoring certain families (i.e., without increasing the between-family component of reproductive variance). Disproportionate representation of certain families could bottleneck genetic diversity, reduce the "portfolio effect" of demographic diversity, and increase risk of artificial selection for traits favored in the captive environment. Although headstart subsamples typically contained more full-sibling families than did wild-hatched subsamples, surprisingly, family evenness was slightly higher in the wild-hatched subsamples than in the headstart subsamples, indicating that reproductive variance was actually lower in the wild than in captivity. By removing individuals for the headstart program, competition could be reduced among wild larvae to increase survival and reduce reproductive variance. In this case, headstart cohorts may be more reflective of natural reproductive variance for the species. In any case, differences in family evenness were slight and non-significant.

When comparing headstart subsamples to overall cohorts, subsamples tended to adequately represent the neutral genetic composition of the overall population, but underrepresented the demographic composition of the overall population. Heterozygosity differed little between subsamples and the total, and on average only 2% of alleles were missed by subsamples. However, a substantial fraction of the families comprising the overall cohort - 37% on average - were absent from headstart subsamples. This suggests that headstart programs, if conducted using the procedures employed by biologists assisting with this study, run relatively little risk of bottlenecking neutral genetic diversity in the populations. However, maintaining this

neutral diversity does not guarantee the maintenance of important adaptive variation. The families that headstart cohorts "left behind" may carry important adaptive differences, such as later timed egg-laying or more concealed egg deposits, from the families present in cattle tanks. I predict that over time, continuing to propagate a fraction of the families through headstarting should significantly shift the demographics and adaptive variation of headstart ponds. This would arise through increasing the survival rate of family lineages possessing adaptive traits that would increase their likelihood of inclusion in the headstarting program, such as early egg-laying or less concealed egg deposits, which might not be otherwise favorable for survival. Thus, the adaptive make-up of the population would increasingly reflect the traits selected for through headstarting over time. On a more positive note, because headstart cohorts tend to capture 60-70% of overall full-sibling families, using headstart individuals to repatriate unoccupied ponds would seem to carry a low risk of demographic or genetic founder effects. These findings suggest that 1) headstarting programs could be beneficial for efforts to repatriate unoccupied ponds and increase population redundancy, but that 2) to increase resiliency and reduce risks, headstarting programs should seek to propagate the widest possible spectrum of variation from each cohort, by representing the greatest possible number of contributing families, for example by collecting eggs and larvae for the headstarting program at multiple intervals throughout the breeding season and from all areas of each pond.

Overall recommendations for conservation management

Ambystoma bishopi is a species in decline and listed as endangered by the U.S. Fish and Wildlife Services since 2009 (USFWS 2015). The continued survival of the species relies on the protection and conservation of breeding and non-breeding habitats and the populations living within. This includes effectively representing the scope of genetic diversity of the species in

management and restoration activities, employing management and conservation activities to many populations to ensure population redundancy, and increasing the size and genetic diversity of populations to bolster population resiliency. Based on the findings of this study, Mayhaw and other potential populations on Georgia are relatively unique and should receive preservation/restoration priority to retain the scope of genetic diversity of the species in future conservation management plans. The species should continue to be monitored and managed on the pond scale as this is the smallest semi-independent population unit. Monitoring of ponds provides the ability to determine the number of breeders, family contributions, and levels of H_E in each pond. However, these ponds group into interconnected metapopulations over longer timescales, which sometimes transcend administrative boundaries, implying the need for a regional perspective on recovery efforts (e.g., restoring connectivity) and setting redundancy and representation targets. Based on this study, increasing pond size and increasing hydrologic permanence may allow ponds to support a larger number of breeders and greater genetic diversity. Additional hydrologic data are needed from ponds on Escribano and elsewhere to further characterize this relationship and monitor this all-important index of habitat suitability. These same findings suggest that future reintroduction efforts should be focused on ponds that have at least 5 hectares of area, slow hydrologic recession rate, and are within 500 m of currently occupied ponds. Results suggest that headstart cohorts could be used to repopulate unoccupied breeding ponds with low risk of flooding ponds with individuals from only a few families. Results from headstart analysis showed that headstart samples miss 37% of total family diversity. Over time, this could cause demographic and genetic shifts in populations as a whole. To avoid this outcome, ponds may need to be surveyed for eggs and larvae multiple times at various

intervals throughout the breeding season if the goal is to collect a maximally diverse population sample for translocation to other ponds.

Table 3.1. AMOVA partition of molecular variation between states, among regions within states, among ponds within regions, among cohorts or breeding seasons within ponds, and among individuals within ponds within cohorts using all sampled ponds. *P*-values were based on 10^4 random permutations of individuals at the level being tested.

Source of Variation	Degrees of Freedom	Molecular Variance	Percent of Variation	*P*
Between states	1	505.0	35.3	<0.001
Among regions within states	4	112.2	7.8	<0.001
Among ponds within regions	33	80.6	5.6	<0.001
Among cohorts within ponds	28	19.7	1.4	<0.001
Among individuals within cohorts	1521	713.9	49.9	
Total	1587	1431.4	100.0	

Table 3.2. Genetic differentiation (Weir and Cockerham's F_{ST}) between pairs of ponds. All cohorts were pooled by pond for this analysis

	Bor.	Dit.	P215	P4	P15	P212	P34	P5	P33	Gar.	Hon	May.	Tor.	EP1	EP46	P52	Res.	Ban.	EP5	P49
Bor.	-																			
Dit.	0.17	-																		
P215	0.21	0.21	-																	
P4	0.17	0.17	0.12	-																
P15	0.18	0.17	0.05	0.08	-															
P212	0.23	0.23	0.18	0.08	0.14	-														
P34	0.19	0.19	0.07	0.09	0.03	0.17	-													
P5	0.18	0.17	0.12	0.02	0.08	0.10	0.09	-												
P33	0.18	0.18	0.06	0.07	0.02	0.14	0.04	0.08	-											
Gar.	0.23	0.23	0.22	0.16	0.17	0.23	0.20	0.16	0.17	-										
Hon	0.12	0.09	0.15	0.11	0.11	0.17	0.12	0.12	0.12	0.16	-									
May.	0.38	0.38	0.36	0.31	0.32	0.38	0.37	0.32	0.33	0.42	0.32	-								
Tor.	0.13	0.10	0.16	0.12	0.12	0.18	0.13	0.13	0.12	0.17	0.03	0.33	-							
EP1	0.16	0.13	0.18	0.14	0.14	0.21	0.16	0.15	0.14	0.21	0.05	0.38	0.06	-						
EP46	0.17	0.14	0.19	0.15	0.15	0.21	0.18	0.15	0.15	0.22	0.05	0.38	0.08	0.11	-					
P52	0.22	0.22	0.18	0.08	0.14	0.15	0.17	0.09	0.14	0.23	0.16	0.38	0.17	0.20	0.21	-				
Res.	0.14	0.11	0.17	0.12	0.12	0.19	0.14	0.13	0.13	0.19	0.03	0.35	0.05	0.07	0.09	0.18	-			
Ban.	0.20	0.17	0.22	0.17	0.18	0.25	0.22	0.18	0.18	0.28	0.10	0.44	0.10	0.15	0.18	0.24	0.13	-		
EP5	0.15	0.13	0.18	0.13	0.14	0.20	0.14	0.14	0.14	0.20	0.03	0.36	0.06	0.09	0.09	0.19	0.06	0.14	-	
P49	0.19	0.18	0.13	0.03	0.08	0.11	0.09	0.03	0.08	0.17	0.12	0.33	0.13	0.15	0.16	0.09	0.13	0.19	0.14	-

	P32	P112	P30	P234	P51	P16	Stan.	Tad.	EP15	EP47	Gho.	P53	P213	Gum	P19	H5	H6	H8
P32	-																	
P112	0.08	-																
P30	0.07	0.02	-															
P234	0.23	0.18	0.19	-														
P51	0.13	0.07	0.06	0.25	-													
P16	0.11	0.11	0.10	0.23	0.13	-												
Stan.	0.17	0.13	0.14	0.17	0.18	0.18	-											
Tad.	0.15	0.08	0.07	0.23	0.18	0.18	0.08	-										
EP15	0.17	0.14	0.14	0.18	0.18	0.18	0.10	0.09	-									
EP47	0.09	0.03	0.05	0.15	0.10	0.12	0.09	0.03	0.09	-								
Gho.	0.12	0.08	0.08	0.13	0.12	0.13	0.06	0.00	0.05	0.04	-							
P53	0.06	0.02	0.00	0.15	0.05	0.09	0.10	0.04	0.10	0.04	0.05	-						
P213	0.09	0.02	0.02	0.17	0.08	0.12	0.13	0.07	0.13	0.02	0.07	0.03	-					
Gum	0.13	0.09	0.09	0.13	0.13	0.14	0.06	0.01	0.04	0.04	0.01	0.06	0.08	-				
P19	0.15	0.03	0.08	0.27	0.18	0.16	0.18	0.26	0.17	-0.01	0.10	0.06	-0.02	0.10	-			
H5	0.10	0.04	0.02	0.16	0.08	0.12	0.10	0.03	0.10	0.03	0.04	0.02	0.01	0.05	0.09	-		
H6	0.06	0.00	0.02	0.16	0.08	0.09	0.11	0.08	0.11	0.01	0.06	0.01	0.02	0.06	0.05	0.04	-	
H8	0.11	0.05	0.06	0.15	0.11	0.13	0.09	0.03	0.08	0.01	0.03	0.04	0.04	0.03	0.05	0.05	0.03	-
H4	0.11	0.02	0.06	0.17	0.11	0.13	0.12	0.09	0.11	0.00	0.06	0.05	0.04	0.06	0.03	0.06	0.01	0.03

Table 3.3. Estimates of effective number of breeders (N_b) for each sampled cohort with at least 10 sampled individuals at each pond on Escribano and Eglin. Means ($N_{b\text{-}mean}$) were obtained by taking the harmonic mean of Colony ($N_{b\text{-}sf}$) and NeEstimator ($N_{b\text{-}LD}$) estimates. Treatment (Tmnt) indicates if ponds were part of the headstart program (HS) or not (W). Year indicates first year of the breeding season. Parentheses contain 5th and 95th confidence intervals for individual pond-cohort estimates and standard deviations across pond-cohort estimates for averages.

Pond	Year	*n*	Tmnt	$N_{b\text{-}SF}$	$N_{b\text{-}LD}$	$N_{b\text{-}mean}$	FS families	*FE*
Borrow	2018	33	HS	31 (19, 55)	22 (14, 36)	26	21	0.953
Borrow	2019	34	HS	29 (18, 53)	20 (12, 36)	24	21	0.937
Borrow	2020	17	HS	25 (14, 48)	8 (3, 16)	12	7	0.948
Ditch	2018	51	HS	7 (4, 21)	11 (8, 14)	9	11	0.785
Ditch	2019	38	HS	5 (2, 20)	4 (3, 4)	4	2	0.968
Ditch	2020	41	HS	7 (4, 21)	8 (5, 11)	7	10	0.794
EP1	2019	11	W	11 (6,28)	9 (3, 31)	10	5	0.804
EP15	2019	67	HS	7 (4, 21)	13 (10, 17)	9	24	0.729
EP46	2019	19	W	9 (5, 24)	9 (7, 11)	9	5	0.795
EP5	2019	11	W	28 (14, 72)	18 (8, 67)	22	8	0.984
Ghost	2018	31	HS	23 (13, 43)	35 (20, 74)	28	20	0.944
Ghost	2019	34	HS	46 (28, 77)	23 (15, 38)	31	28	0.879
Ghost	2020	126	HS	61 (42, 86)	53 (39, 74)	57	63	0.847
Gum	2019	57	HS	42 (27, 66)	34 (23, 53)	37	32	0.923
Gum	2020	80	HS	25 (16, 43)	24 (18, 31)	24	34	0.874
H4	2019	14	W	13 (7, 33)	10 (3, 42)	12	8	0.947
H8	2019	16	W	11 (6, 30)	10 (5, 23)	11	10	0.927
Honey	2018	34	HS	112 (72, 208)	97 (45, 1416)	104	32	0.766
Honey	2019	44	HS	99 (66, 159)	104 (59, 293)	101	38	0.786
Honey	2020	39	HS	20 (12, 39)	72 (37, 300)	31	23	0.852
P112	2020	12	HS	26 (11, 90)	10 (3, 44)	15	9	0.973
P15	2019	19	HS	62 (34, 159)	75 (33, ∞)	68	15	0.969
P15	2020	17	HS	25 (14, 50)	33 (12, ∞)	29	9	0.904
P16	2020	28	HS	10 (5, 26)	14 (9, 22)	12	12	0.871
P212	2019	28	HS	7 (3, 21)	4 (3, 8)	5	5	0.791
P213	2019	15	W	10 (5, 26)	9 (3, 28)	9	9	0.86
P215	2019	26	HS	8 (4, 21)	10 (7, 14)	9	10	0.845
P215	2020	20	HS	5 (2, 20)	3 (2, 3)	4	2	0.811
P33	2018	21	HS	37 (20, 73)	15, (6, 49)	21	14	0.96
P33	2019	38	HS	16 (9, 34)	16 (13, 20)	16	11	0.898
P4	2016	33	HS	33 (21, 58)	34 (22, 62)	34	21	0.94
P4	2017	33	HS	28 (17, 51)	22 (15, 33)	24	15	0.925
P4	2018	40	HS	12 (6, 26)	20 (14, 30)	15	20	0.823
P49	2019	11	HS	46 (20, 151)	25 (11, 629)	36	8	0.971
P5	2016	25	HS	30 (18, 53)	20 (14, 29)	24	11	0.946
P5	2017	32	HS	22 (13, 42)	20 (14, 29)	21	15	0.929
P5	2018	36	HS	28 (18, 50)	33 (24, 47)	30	16	0.935
P52	2019	19	W	9 (5, 24)	6 (3, 12)	7	4	0.893
P53	2019	21	HS	65 (38, 172)	50 (28, 150)	56	19	0.989
P53	2020	25	HS	33 (20, 61)	36 (25, 58)	35	19	0.98
Restoration	2019	11	W	25 (12, 61)	15 (4, 1817)	19	7	0.924
Stanley	2018	20	HS	18 (10, 38)	9 (6, 13)	12	7	0.958
Torpedo	2018	31	HS	44 (27, 73)	53 (35, 99)	48	23	0.967
Torpedo	2019	104	HS	19 (11, 38)	19 (15, 23)	19	32	0.835
Torpedo	2020	33	HS	9 (5, 24)	9 (6, 13)	9	11	0.756
All Escribano	all years				Mean (SD)	28 (26)	20 (14)	
All Eglin	all years				Mean (SD)	23 (16)	12 (5)	
All headstart	2019				Mean (SD)	25 (28)	19 (11)	
All wild	2019				Mean (SD)	12 (7)	7 (2)	

Table 3.4. Genomic diversity by *A. bishopi* breeding pond. Cohorts were pooled by pond for this analysis. Statistics include number of individuals (*n*), observed heterozygosity (H_O), and expected heterozygosity (H_E).Heterozygosity values are means across loci (standard deviations in parentheses) Treatment (Tmnt) indicates if ponds were part of the headstart program (HS) or not (W).

Pond	Region	*n*	Tmnt	H_O	H_E
Borrow	Escribano	84	HS	0.27 (0.19)	0.28 (0.18)
Ditch	Escribano	130	HS	0.28 (0.19)	0.28 (0.18)
EP15	Escribano	67	HS	0.32 (0.20)	0.29 (0.17)
Stanley	Escribano	20	HS	0.31 (0.21)	0.30 (0.18)
Honey	Escribano	117	HS	0.31 (0.16)	0.32 (0.15)
Gum	Escribano	137	HS	0.30 (0.15)	0.32 (0.15)
Ghost	Escribano	191	HS	0.31 (0.15)	0.32 (0.14)
EP46	Escribano	19	W	0.32 (0.24)	0.29 (0.19)
EP1	Escribano	11	W	0.33 (0.24)	0.30 (0.19)
EP5	Escribano	11	W	0.30 (0.21)	0.30 (0.18)
EP47	Escribano	9	W	0.29 (0.21)	0.32 (0.18)
Restoration	Escribano	11	W	0.32 (0.22)	0.31 (0.18)
Torpedo	Escribano	168	HS	0.31 (0.16)	0.31 (0.16)
P215	Eastbay	49	HS	0.30 (0.20)	0.29 (0.18)
P33	Eastbay	59	HS	0.30 (0.18)	0.31 (0.16)
P15	Eastbay	39	HS	0.35 (0.19)	0.32 (0.16)
P112	Eastbay	14	HS	0.29 (0.19)	0.32 (0.17)
P16	Eastbay	28	HS	0.29 (0.20)	0.29 (0.18)
P32	Eastbay	8	W	0.29 (0.23)	0.30 (0.20)
P212	Oglesby	31	HS	0.30 (0.22)	0.28 (0.19)
P5	Oglesby	95	HS	0.31 (0.16)	0.31 (0.16)
P4	Oglesby	118	HS	0.32 (0.16)	0.32 (0.16)
P53	Oglesby	46	HS	0.32 (0.15)	0.34 (0.14)
P52	Oglesby	19	W	0.33 (0.24)	0.28 (0.19)
P213	Oglesby	15	W	0.29 (0.19)	0.32 (0.17)
H4	Hurlburt	14	W	0.34 (0.21)	0.32 (0.17)
H8	Hurlburt	16	W	0.31 (0.19)	0.33 (0.16)
H5	Hurlburt	9	W	0.32 (0.21)	0.33 (0.17)
H6	Hurlburt	8	W	0.36 (0.25)	0.33 (0.19)
All headstart			Mean (SD)	0.31 (0.02)	0.31 (0.02)
All wild			Mean (SD)	0.32 (0.02)	0.31 (0.02)

Table 3.5. Diversity statistics for study regions. Entries are means (standard deviation in parentheses) across all ponds in a region. I excluded ponds with sample size less than 4 and ponds with hybrid origin. Allelic richness (A_4) is the average number of alleles per locus per four individuals.

Region	Ponds	Individuals sampled per pond	Polymorphic loci	A_4	H_O	H_E
Garcon	1	4	3351	1.29	0.37	0.29
Escribano	14	70 (66.9)	4269.6 (444.4)	1.30 (0.02)	0.31 (0.02)	0.30 (0.02)
Eastbay	8	25.8 (21.3)	3990.1 (568.4)	1.30 (0.02)	0.30 (0.02)	0.30 (0.02)
Hurlburt	4	11.8 (3.4)	4287.5 (110.3)	1.33 (0.01)	0.32 (0.03)	0.33 (0.01)
Oglesby	7	46.9 (43.3)	4209.3 (512.5)	1.31 (0.02)	0.31 (0.01)	0.31 (0.02)
Mayhaw	1	4	2395	1.22	0.25	0.22

Table 3.6. Genomic diversity statistics for headstart subsamples (collected as eggs, hatched in cattle tanks), wild subsamples (hatched in wild, sampled as juveniles), and overall total cohorts (headstart and wild combined), for ponds and breeding years that were part of the headstarting program at Escribano and Eglin.

		Headstart					Wild					Total				
Pond	Year	*n*	Private alleles	$N_{b\text{-}mean}$	Unique families	*FE*	*n*	Private alleles	$N_{b\text{-}mean}$	Unique families	*FE*	*n*	Total alleles	$N_{b\text{-}mean}$	Total families	*FE*
Ditch	2018	42	334	8	6	0.81	9	218	13	3	0.94	51	8820	9	11	0.79
Borrow	2019	23	143	23	13	0.94	11	529	22	5	0.95	34	9067	24	21	0.94
Ditch	2019	20	34	5	0	0.97	18	176	5	0	0.96	38	8180	4	2	0.97
EP15	2019	59	578	8	17	0.70	8	66	97	6	0.98	67	9366	9	24	0.73
Gum	2019	43	487	35	26	0.85	14	43	5	6	0.88	57	9381	37	32	0.92
Torpedo	2019	59	31	9	4	0.85	45	208	36	24	0.90	104	9358	19	32	0.84
Borrow	2020	10	992	11	5	0.94	7	145	4	2	0.99	17	8817	12	7	0.95
Ditch	2020	31	429	7	3	0.78	10	36	9	3	0.97	41	8851	7	10	0.79
Ghost	2020	46	9	21	9	0.93	80	231	6	51	0.39	126	9498	57	63	0.85
Gum	2020	53	73	17	20	0.85	27	81	14	13	0.88	80	9496	24	34	0.87
Honey	2020	9	39	2	6	1.00	30	458	24	17	0.84	39	9359	31	23	0.85
Torpedo	2020	22	231	6	4	0.62	11	301	19	5	0.95	33	9060	9	11	0.76
P4	2016	23	291	27	12	0.92	10	115	61	5	0.99	33	9222	34	21	0.94
P5	2016	15	364	26	4	0.95	10	88	41	2	0.97	25	9007	24	11	0.95
P4	2017	23	724	26	12	0.95	10	63	5	3	0.87	33	9206	24	15	0.93
P15	2019	8	272	1	3	1.00	11	259	110	9	1.00	19	9118	68	15	0.97
P215	2019	10	283	23	3	0.97	16	217	8	4	0.85	26	8842	9	10	0.85
Average		29	313	15	9	0.88	19	190	28	9	0.90	48	9097	24	20	0.87

.

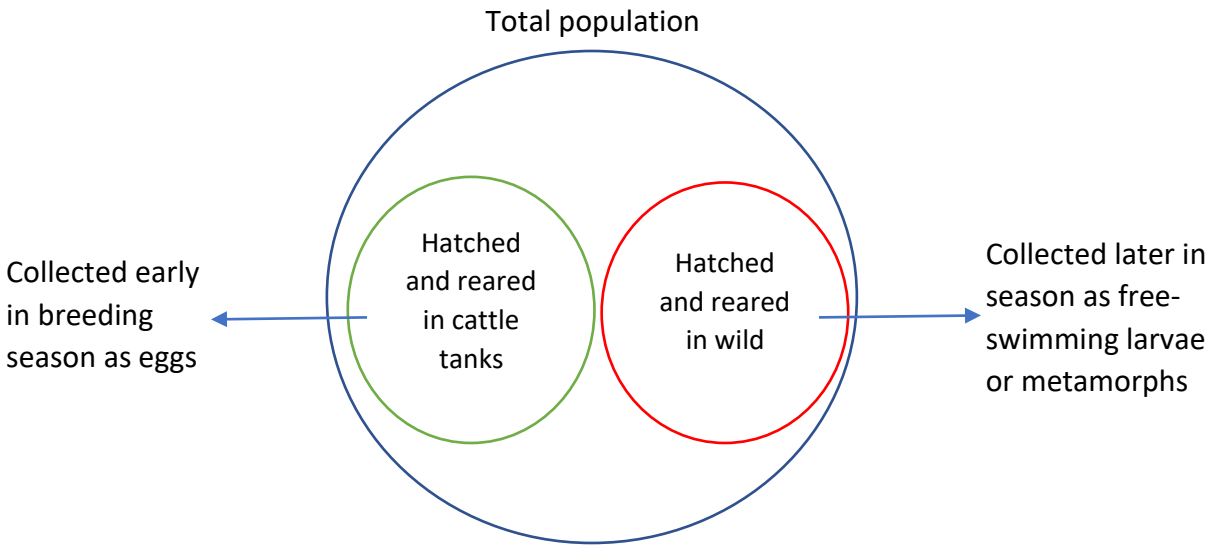

Figure 3.1. Visualization of the two groups for comparison in assessing genetic representation of headstarting cohorts. Genetic diversity and demographic estimates for individuals from a breeding pond that are collected as eggs and hatched in cattle tanks (a) were compared to those of the entire pond population (b) which includes individuals hatched in cattle tanks as well as those found as larvae later in the breeding season

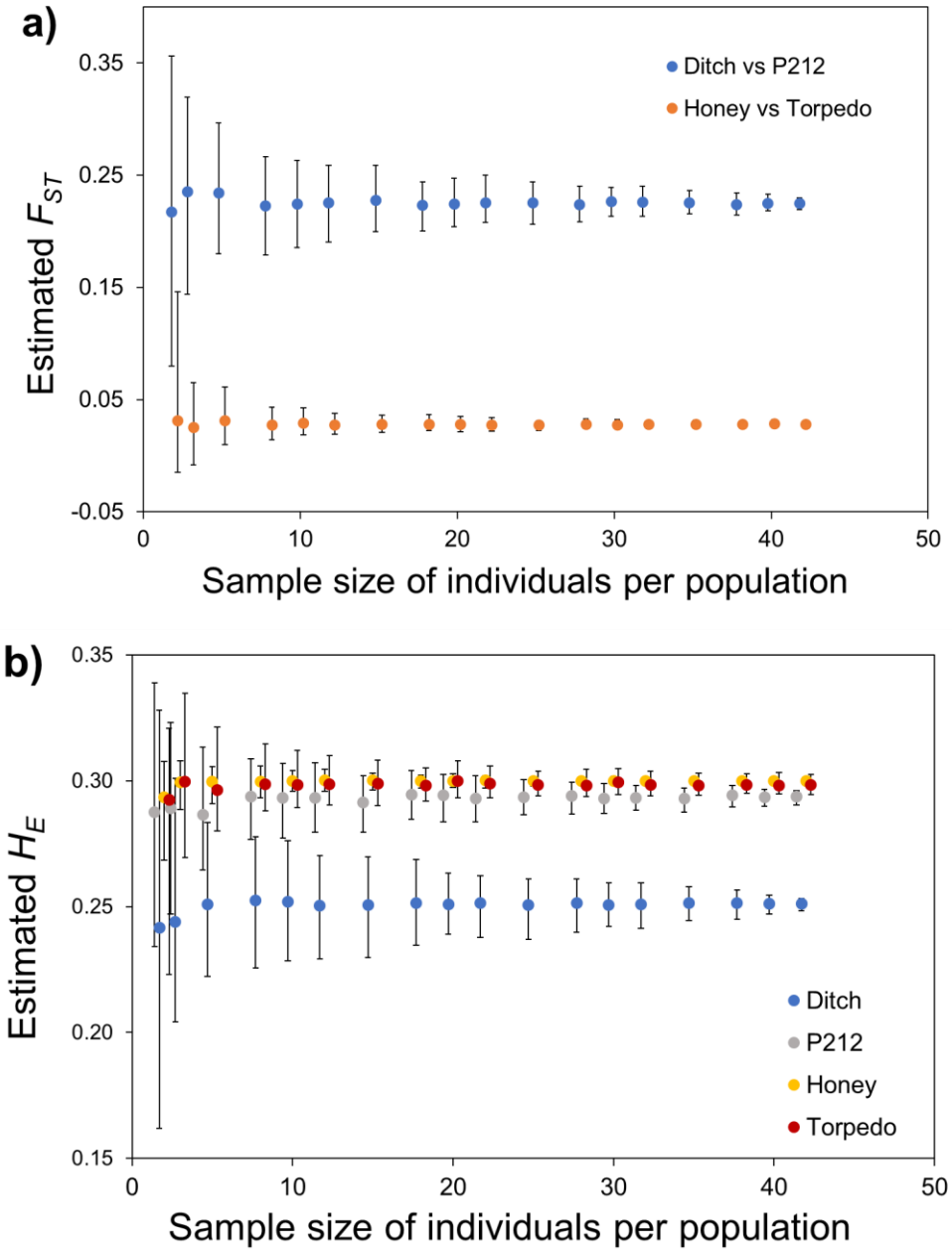

Figure 3.2. Results of rarefaction analyses for a) F_{ST} and b) H_E using four ponds to determine minimum sample size requirements for downstream analyses. Points represent means and error bars represent 5th and 95th percentiles based on 1000 randomly selected loci at each level of sampling intensity.

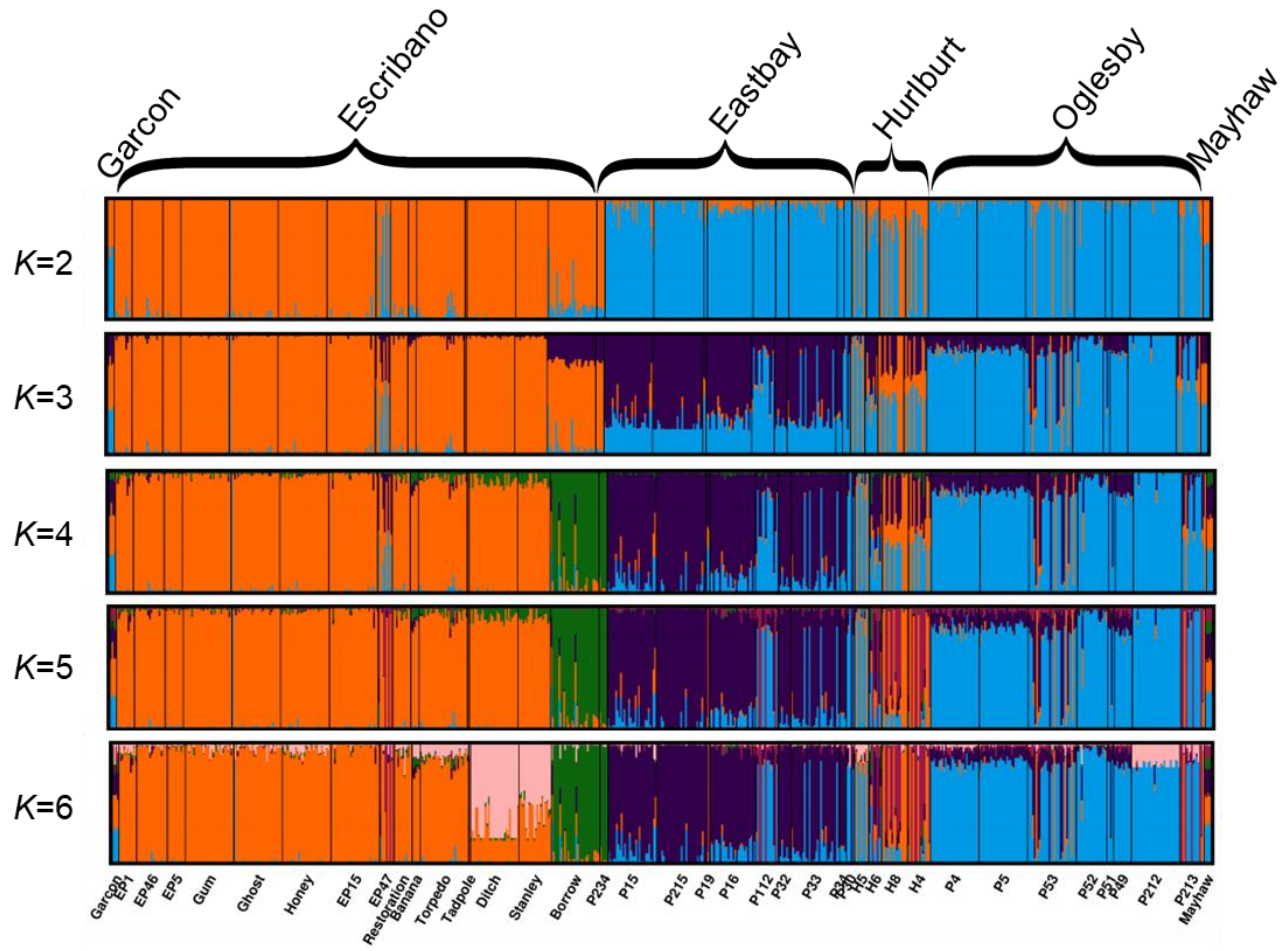

Figure 3.3. Range-wide STRUCTURE bar plot for *K* values 2-6 using individuals from Garcon, Escribano, Eglin, and Mayhaw (*n*=678). Ponds were limited to 30 random individuals.

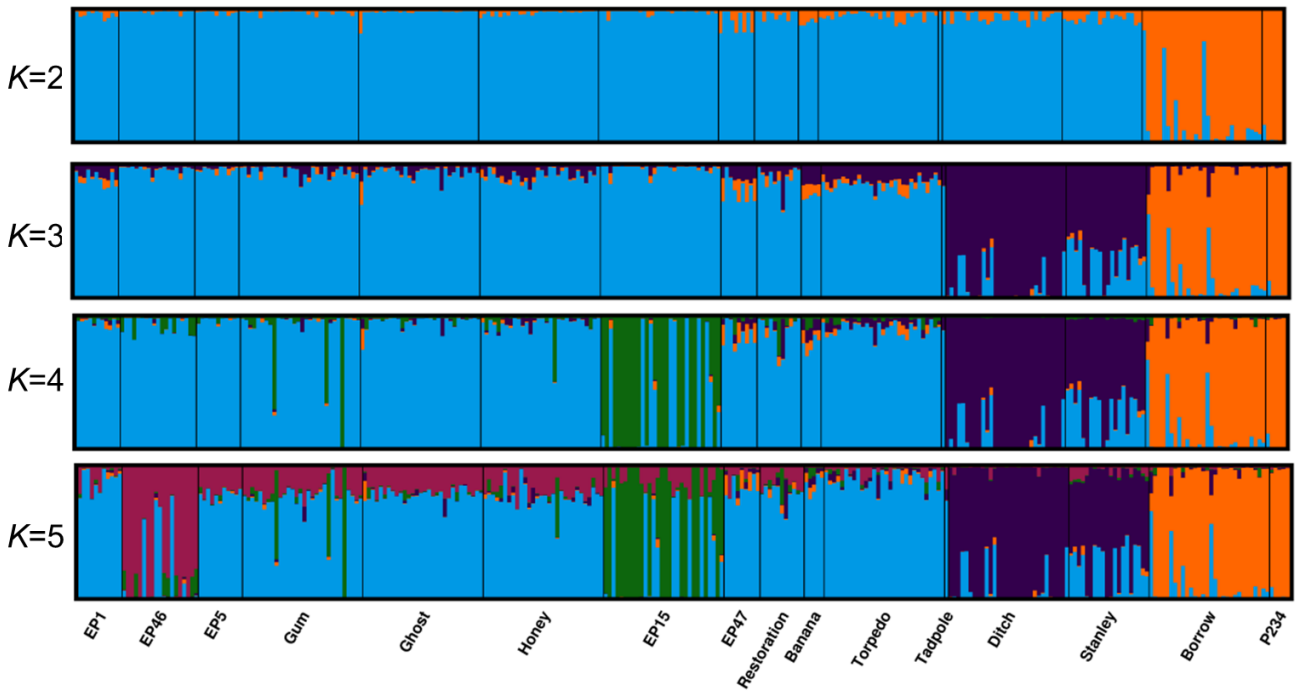

Figure 3.4. STRUCTURE bar plot results for *K* 2-5 for ponds on Escribano (*n*=302). Ponds were limited to 30 randomly selected individuals.

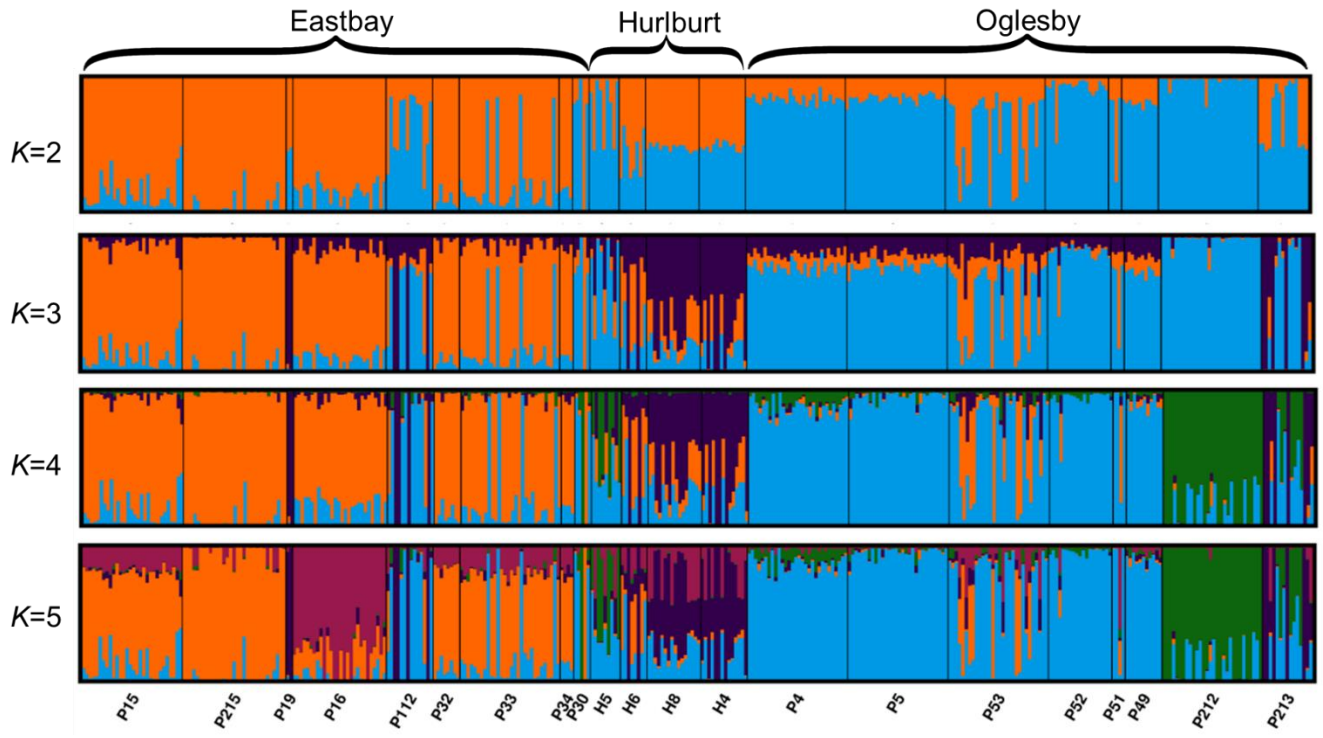

Figure 3.5. STRUCTURE bar plot results for *K* 2-5 for ponds on Eglin (*n*=368). Ponds were limited to 30 randomly selected individuals.

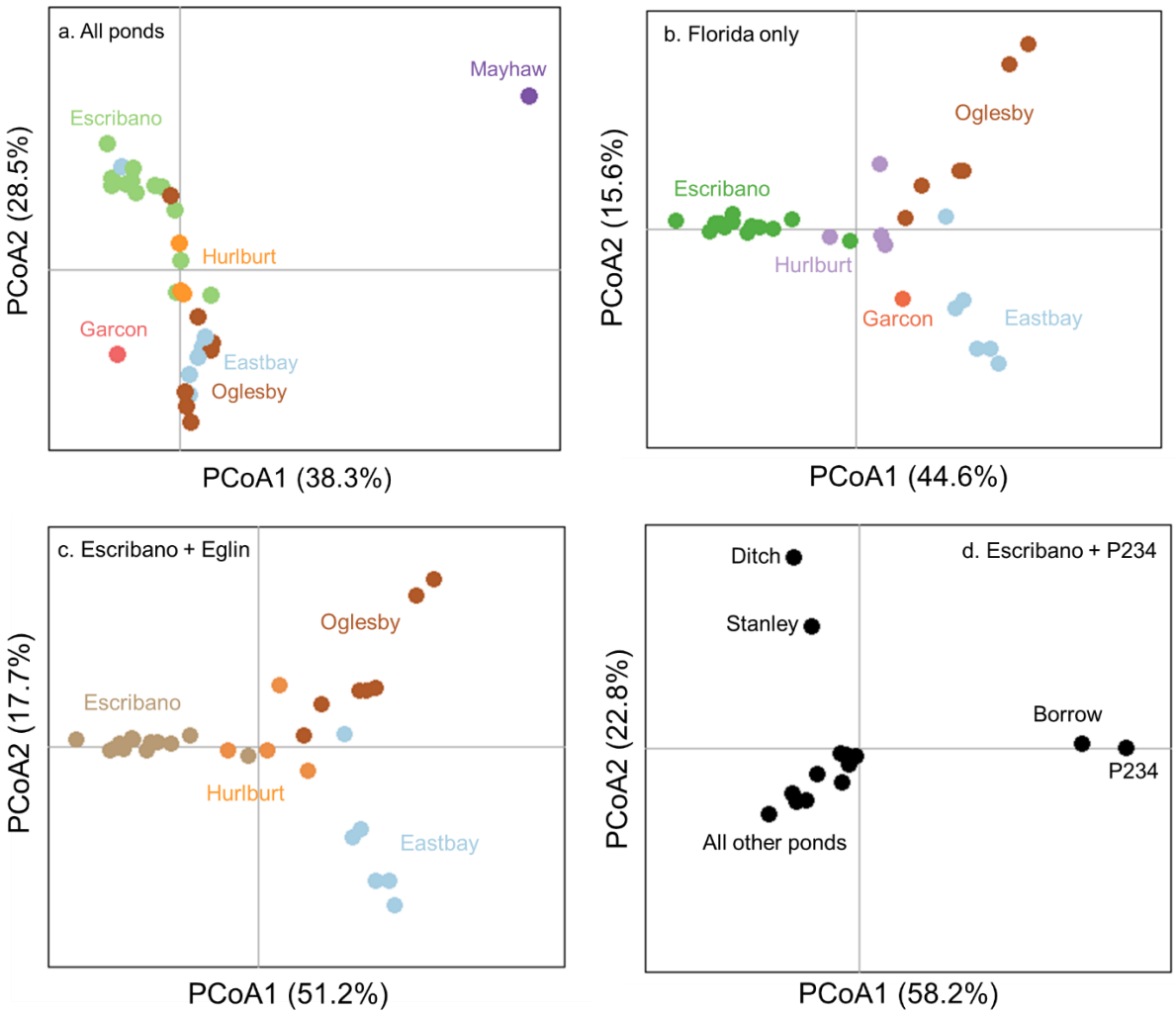

Figure 3.6. Principal coordinates analysis (PCoA) results based on inclusion of a) all ponds, b) ponds in Florida only (i.e., excluding Mayhaw), c) ponds on Escribano and Eglin only, d) ponds on Escribano plus Eglin pond 234.

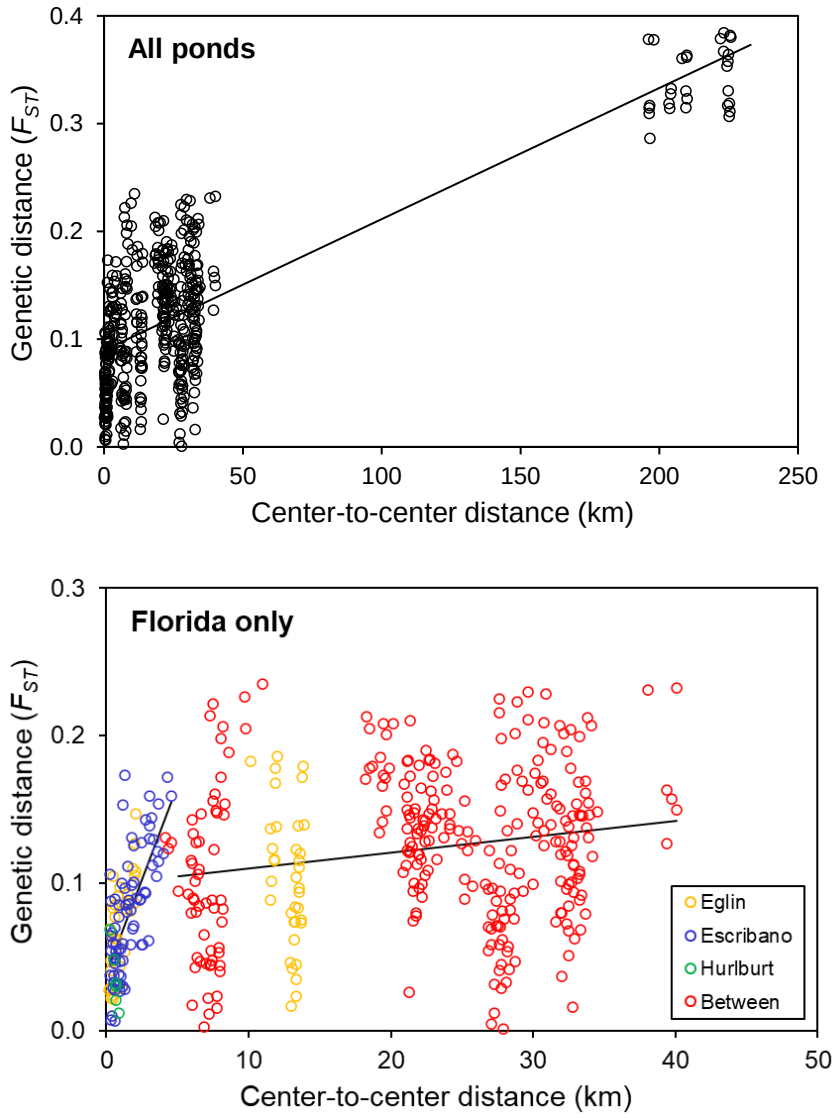

Figure 3.7. Comparison of pairwise genetic distance (F_{ST}) to center-to-center Euclidean distance (km) between ponds. Linear trend lines are fit to relationships based on multiple regression on distance matrices results. Graphs include all ponds (F_{ST} =0.001(km) + 0.0897 ; R^2=0.61 ; p=0.0001), and ponds located in Florida. For the latter, the trend is separately shown for comparisons less than 5 km apart (F_{ST} =0.025(km) + 0.0398 ; R^2=0.50 ; p=0.0001) vs. greater than 5 km apart (F_{ST} =0.001(km) + 0.0991 ; R^2=0.04 ; p=0.0242). Points are colored to indicate pairwise comparisons between properties (red) vs. within properties (blue, yellow, green).

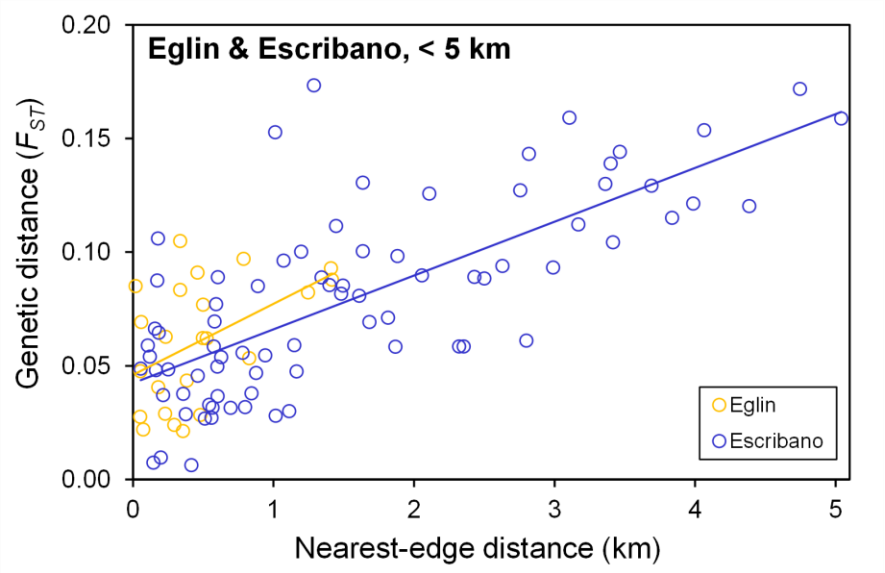

Figure 3.8. Comparison of pairwise genetic distances (F_{ST}) to pond nearest-edge Euclidian distance (km) for ponds on Eglin (yellow, F_{ST} =0.036(km) + 0.0442 ; R^2=0.53 ; p=0.0008) and Escribano (blue, F_{ST} =0.024(km) + 0.0423 ; R^2=0.51 ; p=0.0001) within 5 km of each other.

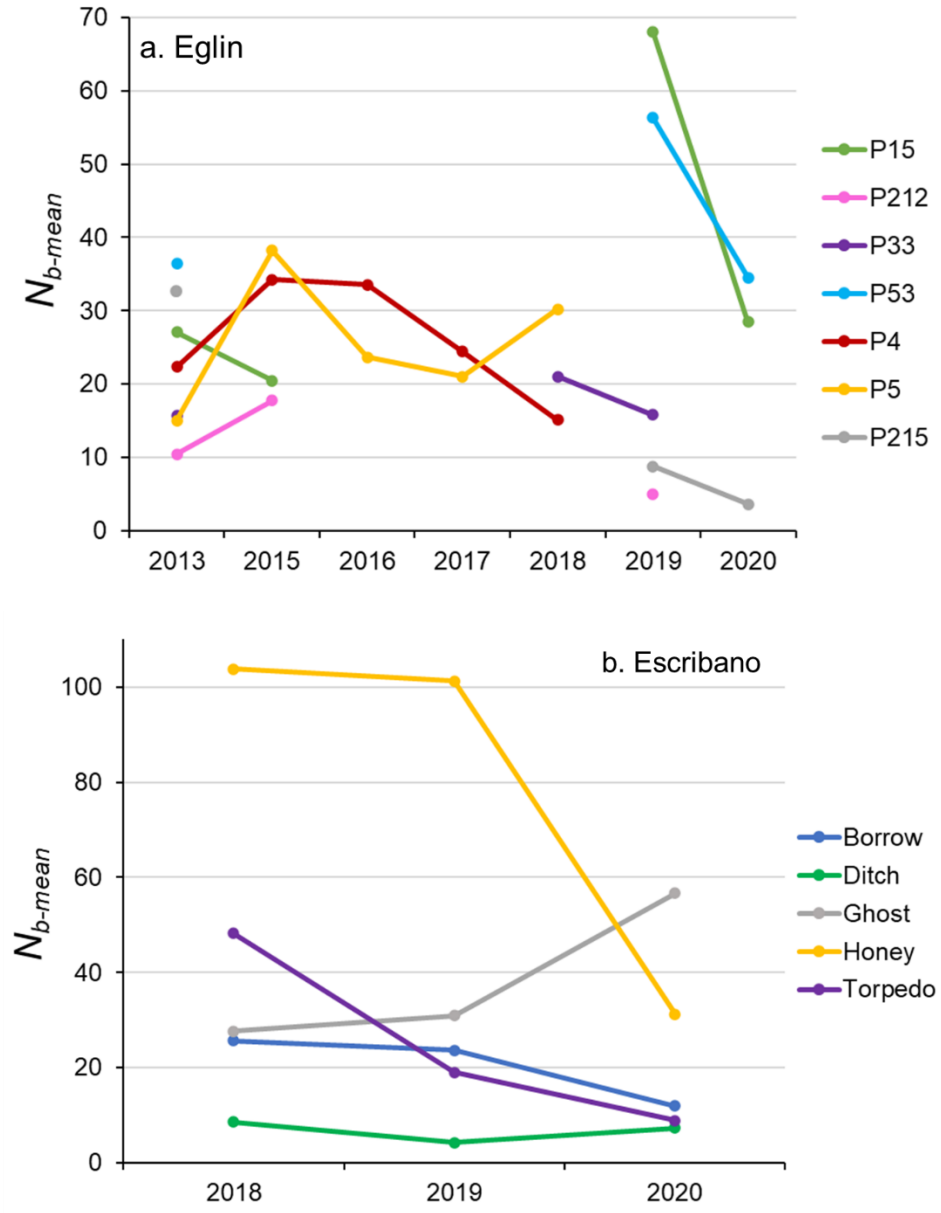

Figure 3.9. Estimated mean effective number of breeders for ponds and breeding seasons in which $N_{b\text{-}mean}$ was estimable for multiple years. Estimates were obtained by taking the harmonic mean of COLONY and NeEstimator results for ponds with ≥10 individuals. Estimates from 2013 and 2015 were obtained by reanalyzing Wendt et al.'s microsatellite data using settings applied on SNP data.

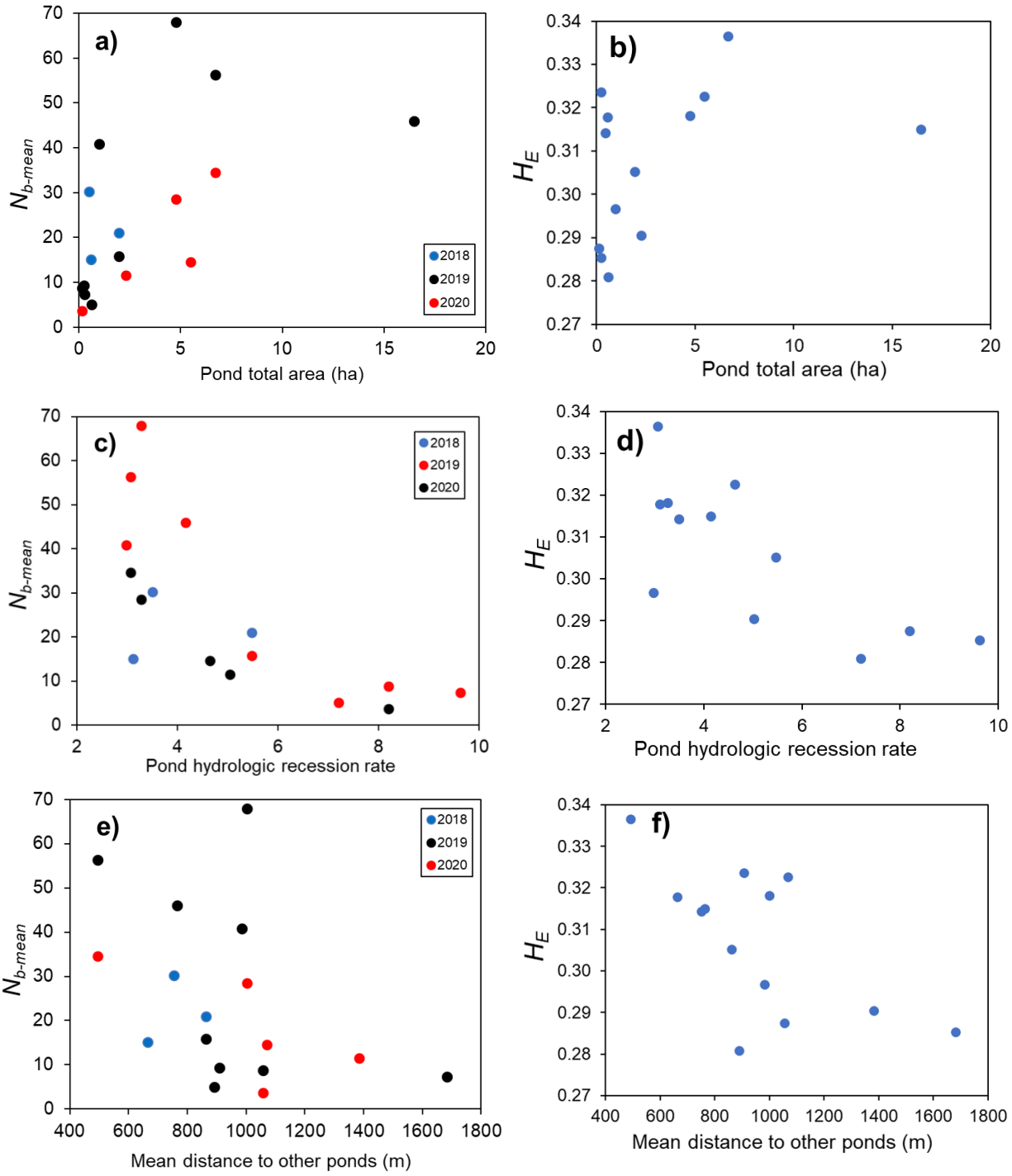

Figure 3.10. Relationships of effective number of breeders ($N_{b\text{-}mean}$; panels a, c, e) and expected heterozygosity (H_E; panels b, d, f) with pond total area ($N_{b\text{-}mean}$ r =0.58, p=0.014; H_E r =0.41,

p=0.16), pond hydrologic recession rate ($N_{b\text{-}mean}$ r = -0.71, p=0.002; H_E r =-0.75, p=0.003), and mean distance to other ponds ($N_{b\text{-}mean}$ r = -0.44, p=0.078; H_E r =-0.65, p=0.017) for ponds on the Oglesby and Eastbay flatwoods regions of Eglin.

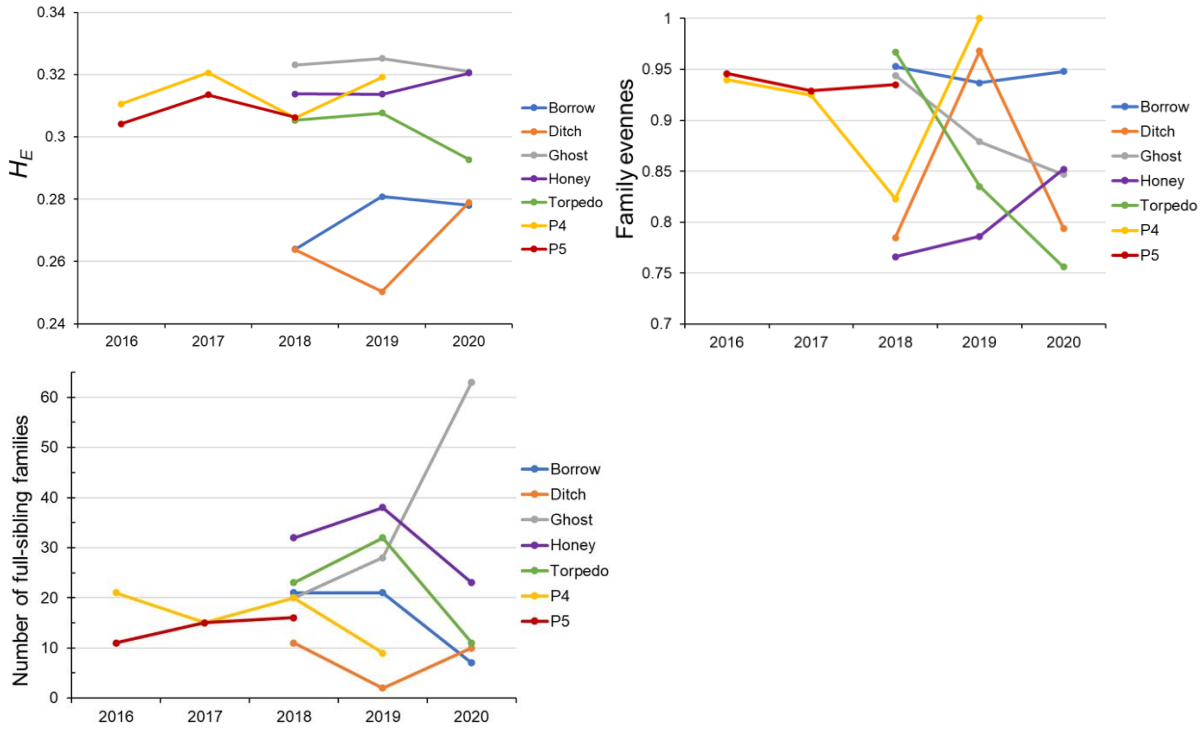

Figure 3.11. Expected heterozygosity (H_E), family evenness, and number of full-sibling families over time for pond-cohorts with ≥8 sampled individuals at a) Escribano and b) Eglin. All ponds-cohorts represented are used in the headstart programs.

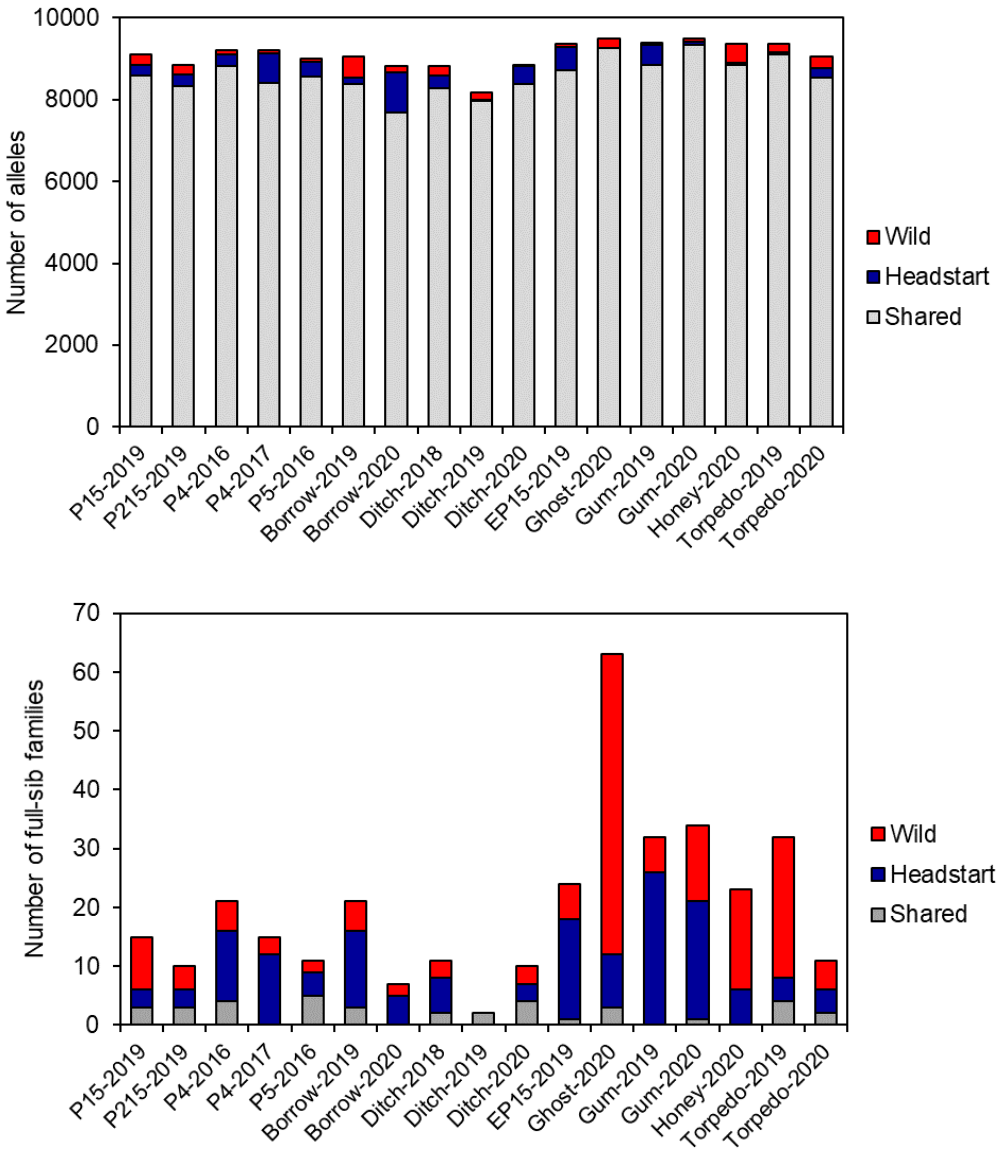

Figure 3.12. Total counts of the number of alleles (top panel) and the number of full-sibling families (bottom panel) that were unique to wild subsamples (red), unique to headstart subsamples (blue), or shared between wild and headstart subsamples (gray). The total height of each bar indicates the count for the total cohort, whereas the red fraction indicates the fraction that was "missed" by headstart subsamples. Results are shown only for cohorts in which both headstart and wild subsamples consisted of at least seven individuals.

REFERENCES

Ackerman, M. W., Hand, B. K., Waples, R. K., Luikart, G., Waples, R. S., Steele, C. A., ... & Campbell, M. R. (2017). Effective number of breeders from sibship reconstruction: empirical evaluations using hatchery steelhead. *Evolutionary applications*, *10*(2), 146-160.

Alford, R. A. (2010). Declines and the global status of amphibians. *Ecotoxicology of amphibians and reptiles*, *2*, 13-46.

Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics*, *202*(2), 389-400.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature reviews genetics*, *11*(10), 697-709.

Amburgey, S., Funk, W. C., Murphy, M., & Muths, E. (2012). Effects of hydroperiod duration on survival, developmental rate, and size at metamorphosis in boreal chorus frog tadpoles (Pseudacris maculata). *Herpetologica*, *68*(4), 456-467.

Anderson, J. D., & Williamson, G. K. (1976). Terrestrial mode of reproduction in Ambystoma cingulatum. *Herpetologica*, 214-221.

Araki, H., Cooper, B., & Blouin, M. S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science*, *318*(5847), 100-103.

Arnfield, H., Grant, R., Monk, C., & Uller, T. (2012). Factors influencing the timing of spring migration in common toads (B ufo bufo). *Journal of Zoology*, *288*(2), 112-118.

Balkenhol, N., Cushman, S., Storfer, A., & Waits, L. (2015). Landscape genetics: concepts, methods, applications. John Wiley & Sons.

Balkenhol, N., Dudaniec, R. Y., Krutovsky, K. V., Johnson, J. S., Cairns, D. M., Segelbacher, G., ... & Joost, S. (2017). Landscape genomics: understanding relationships between environmental heterogeneity and genomic characteristics of populations. In *Population genomics* (pp. 261-322). Springer, Cham.

Banks, B., Denton, J. S., & Beebee, T. J. (1993). Long-term management of a natterjack toad (Bufo calamita) population in southern Britain. *Amphibia-Reptilia*, *14*(2), 155-168.

Barata, I. M., Uhlig, V. M., Silva, G. H., & Ferreira, G. B. (2016). Downscaling the gap: protected areas, scientific knowledge and the conservation of amphibian species in Minas Gerais, southeastern Brazil. *South American Journal of Herpetology*, *11*(1), 34-45.

Bartlett R, Bartlett P (2006) Guide and reference to the amphibians of eastern and central North America (north of Mexico). University Press of Florida, Gainesville.

Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., Bentley, K. E., Finger, J. W., Louha, S., Troendle, N., Diaz-Jaimes, P., Mauricio, R., & Faircloth, B. C. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). PeerJ, 7, e7724. https://doi.org/10.7717/peerj.7724

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M., Beacham, T. D., ... & Withler, R. E. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology & Evolution*, *32*(9), 665-680.

Bevelhimer M, Stevenson D, Giffen N, Ravenscroft K (2008) Annual Surverys of Larval Ambystoma cingulatum Reveal Large Differences in Dates of Pond Residency. Southern Naturalist, 7(2), 311-322.

Bishop D, Haas C (2005) Burning trends and potential negative effects of suppressing wetland fires on flatwoods salamanders. Natural Areas Journal, 25, 290-294.

Blaustein A, Han BA, Relyea RA, Johnson PT, Buck JC, Gervasi SS, Kats LB (2011) The complexity of amphibian population declines: understanding the role of cofactors in driving amphibian losses. Annuals of the New York Academy of Science, 1223, 108-119.

Blouin, M. S., Phillipsen, I. C., & Monsen, K. J. (2010). Population structure and conservation genetics of the Oregon spotted frog, Rana pretiosa. *Conservation Genetics*, *11*(6), 2179-2194.

Bohenek, J. R., & Resetarits, W. J. (2018). Are direct density cues, not resource competition, driving life history trajectories in a polyphenic salamander?. *Evolutionary Ecology*, *32*(4), 335-357.

Bonin, A., Bellemain, E., Eidesen, P. B., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. Molecular Ecology, 13(11), 3261–3273. https://doi.org/10.1111/j.1365-294X.2004.02346.x

Brooks, G. C., Gorman, T. A., Jiao, Y., & Haas, C. A. (2020). Reconciling larval and adult sampling methods to model growth across life-stages. *PLOS ONE*, *15*(8), e0237737. https://doi.org/10.1371/journal.pone.0237737

Brooks, G. C., Smith, J. A., Frimpong, E. A., Gorman, T. A., Chandler, H. C., & Haas, C. A. (2019). Indirect connectivity estimates of amphibian breeding wetlands from spatially

explicit occupancy models. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *29*(11), 1815–1825. https://doi.org/10.1002/aqc.3190

Brooks, G. C., Smith, J. A., Gorman, T. A., & Haas, C. A. (2019). Discerning the Environmental Drivers of Annual Migrations in an Endangered Amphibian. *Copeia*, *107*(2), 270. https://doi.org/10.1643/ch-18-068

Buckley L, Jetz W (2007) Environmental and historical constraints on global patterns of amphibian richness. Proceedings of the Royal Society B, 274, 1167-1173.

Burton T, Likens G (1975a) Salamander populations and biomass in the Hubbard Brook experimental forest, New Hampshire. Copeia, 1975, 541-546.

Cabrera-Guzmán, E., Crossland, M. R., Brown, G. P., & Shine, R. (2013). Larger body size at metamorphosis enhances survival, growth and performance of young cane toads (Rhinella marina). *PloS one*, *8*(7), e70121.

Camacho-Sanchez, M., Velo-Antón, G., Hanson, J. O., Veríssimo, A., Martínez-Solano, Í., Marques, A., ... & Carvalho, S. B. (2020). Comparative assessment of range-wide patterns of genetic diversity and structure with SNPs and microsatellites: A case study with Iberian amphibians. *Ecology and evolution*, *10*(19), 10353-10363.

Cammen, K. M., Schultz, T. F., Don Bowen, W., Hammill, M. O., Puryear, W. B., Runstadler, J., ... & Kinnison, M. (2018). Genomic signatures of population bottleneck and recovery in Northwest Atlantic pinnipeds. *Ecology and Evolution*, *8*(13), 6599-6614.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular ecology*, *22*(11), 3124-3140.

Catchen, J., Rochette, N., Rivera-Colon, A., Cresko, W., Hohenlohe, P., Amores, A., Bassham, S., & Postlethwait, J. (2013) Stacks Manual

Chandler, H. C., Haas, C. A., & Gorman, T. A. (2015). The effects of habitat structure on winter aquatic invertebrate and amphibian communities in pine flatwoods wetlands. *Wetlands*, *35*(6), 1201-1211.

Chandler, H. C., Rypel, A. L., Jiao, Y., Haas, C. A., & Gorman, T. A. (2016). Hindcasting Historical Breeding Conditions for an Endangered Salamander in Ephemeral Wetlands of the Southeastern USA: Implications of Climate Change. *PLOS ONE*, *11*(2), e0150169. https://doi.org/10.1371/journal.pone.0150169

Chandler, H. C., McLaughlin, D. L., Gorman, T. A., McGuire, K. J., Feaga, J. B., & Haas, C. A. (2017). Drying rates of ephemeral wetlands: implications for breeding amphibians. *Wetlands*, *37*(3), 545-557.

Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. *Molecular Ecology*, **18**, 2746– 2765.

Charlesworth B. 2009 Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10, 195– 205. (doi:10.1038/nrg2526)

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nature Review Genetics, 10, 783-796.

Clipp H, Anderson J (2014) Environmental and anthropogenic factors influencing salamanders in riparian forests: A review. Forests, 5, 2679-2702.

Collins, J. P., & Storfer, A. (2003). Global amphibian declines: sorting the hypotheses. *Diversity and distributions*, *9*(2), 89-98.

Collins J (2010) Amphibian decline and extinction: what we know and what we need to learn. Diseases of Aquatic Organisms, 92, 93-99.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK. 2000. Considering evolutionary processes in conservation biology. *Trends Ecol. Evol.* 15:290–95

Cushman S (2006) Effects of habitat loss and fragmentation on amphibians: a review and prospectus. Biological Conservation, 128, 231-240.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.

Daszak P, Cunnigham A, Hyatt A (2003) Infectious disease and amphibian population declines. Diversity and Distribution, 9, 141-150.

Daszak P, Scott DE, Kilpatrick AM, Faggioni C, Gibbons JW, Porter D (2005) Amphibian population declines at Savannah River site are linked to climate, not chytridiomycosis. Ecology, 86, 3232-3237.

Davidson, C (2004) Declining downwind: amphibian population declines in California and historical pesticide use. Ecological Applications, 14, 1892-1902.

de Guia, A. P. O., & Saitoh, T. (2007). The gap between the concept and definitions in the evolutionarily significant unit: the need to integrate neutral genetic variation and adaptive variation. *Ecological Research*, *22*(4), 604-612.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., & Vigouroux, Y. (2013). Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular ecology*, *22*(5), 1383-1399.

DiLeo, M. F., & Wagner, H. H. (2016). A Landscape Ecologist's Agenda for Landscape Genetics. *Current Landscape Ecology Reports*, *1*(3), 115–126. https://doi.org/10.1007/s40823-016-0013-x

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular ecology resources*, *14*(1), 209-214.

Dodd CK, Barichivich WJ, Smith LL (2004). Effectiveness of a barrier wall and culverts in reducing wildlife mortality on a heavily traveled highway in Florida. *Biol Conserv* **118**: 619–631.

Dresser, C. M., Pierson, T. W., & Fitzpatrick, B. M. (2018). Isolation by distance, local adaptation, and fortuitous coincidence of geo-political boundaries with spatial-genetic clusters in southern Bog Turtles. *Global ecology and conservation*, *16*, e00474.

Driscoll DA (1997) Mobility and metapopulation structure of Geocrinia alba and Geogrinia vitelli, two endangered frog species from southwestern Australia. Australian Journal of Ecology, 22, 185-195.

Duellman W, Trueb L (1994) The biology of amphibians. Baltimore: Johns Hopkins University Press.

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

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.

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics*, *1*, 117693430500100003.

Fahrig, L., & Merriam, G. (1994). Conservation of fragmented populations. *Conservation biology*, *8*(1), 50-59.

Fahrig, L., Pedlar, J. H., Pope, S. E., Taylor, P. D., & Wegner, J. F. (1995). Effect of road traffic on amphibian density. *Biological conservation*, *73*(3), 177-182.

Farmer AL, Walls SC, Haas C et al (2016) A statewide species and habitat assessment for the reticulated flatwoods salamander, frosted flatwoods salamander, and striped newt. Florida Fish and Wildlife Annual Report.

Faubet P, Waples RS, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. Molecular Ecology, 16(6), 1149-1166.

Ferchaud, A. L., Perrier, C., April, J., Hernandez, C., Dionne, M., & Bernatchez, L. (2016). Making sense of the relationships between Ne, Nb and Nc towards defining conservation thresholds in Atlantic salmon (Salmo salar). *Heredity*, *117*(4), 268-278.

Fisher RA (1930) The genetical theory of natural selection. Oxford University Press, Oxford

Fisher, R. N., & Shaffer, H. B. (1996). The decline of amphibians in California's Great Central Valley. *Conservation biology*, *10*(5), 1387-1397.

Fischer, M., & Matthies, D. (1998). Effects of population size on performance in the rare plant Gentianella germanica. *Journal of Ecology*, *86*(2), 195-204.

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., ... & Widmer, A. (2017). Estimating genomic diversity and population differentiation–an empirical comparison of microsatellite and SNP variation in Arabidopsis halleri. *BMC genomics*, *18*(1), 1-15Flesch, E. P., Graves, T. A., Thomson, J. M., Proffitt, K. M., White, P. J., Stephenson, T. R., & Garrott, R. A. (2020). Evaluating wildlife translocations using genomics: A bighorn sheep case study. *Ecology and Evolution*.

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, *180*(2), 977-993.

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular ecology*, *25*(1), 104-120.

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2017). Using genotype-environment associations to identify multilocus local adaptation. *bioRxiv*, 129460.

Frankham R. 2005. Genetics and extinction. *Biol. Conserv*. 126:131–40 34.

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R. & Fenster, C.B. (2011). Predicting the probability of outbreeding depression. Conserv. Biol. 25, 465–475.

Frankham R (2009) Effective population size/adult population size ratios in wildlife: a review. Genetic Resources, 66, 95-107.

Franklin, I. R. (1980). Evolutionary change in small populations.

Franklin, I. R., & Frankham, R. (1998, February). How large must populations be to retain evolutionary potential?. In *Animal Conservation forum* (Vol. 1, No. 1, pp. 69-70). Cambridge University Press.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in ecology & evolution*, *27*(9), 489-496.

Funk, W. C., Murphy, M. A., Hoke, K. L., Muths, E., Amburgey, S. M., Lemmon, E. M., & Lemmon, A. R. (2016). Elevational speciation in action? Restricted gene flow associated with adaptive divergence across an altitudinal gradient. *Journal of Evolutionary Biology*, *29*(2), 241-252.

Gamble, L. R., McGarigal, K., & Compton, B. W. (2007). Fidelity and dispersal in the pond-breeding amphibian, Ambystoma opacum: implications for spatio-temporal population dynamics and conservation. *Biological Conservation*, *139*(3-4), 247-257.

Gibbs J (2000) Wetland loss and biodiversity conservation. Conservation Biology, 14, 314-317.

Gibbons J, Winne CT, Scott DE, Willson JD, Glaudas X, Andrews KM, Todd BD, Fedewa LA, Wilkinson L, Tsaliagos RN, Harper SJ, Greene JL, Tuberville TD, Metts BS, Dorvas ME, Nestor JP, Young CA, Akre T, Reed RN, Buhlmann KA, Norman J, Croshaw DA, Hagen C, Rothermel BB (2006) Remarkable amphibian biomass and abundance in an isolated wetland: implications for wetland conservation. Conservation Biology, 20, 1457-1465.

Giery, S. T., Drake, D. L., & Urban, M. C. (2021). Microgeographic evolution of metabolic physiology in a salamander metapopulation. *Ecology*, *102*(11), e03488.

Gill DE (1978) The metapopulation ecology of the red spotted newt, Notophthalmus viridescens (Rafinesque). Ecological Monographs, 48, 145-166.

Golladay, S. W., Clayton, B. A., Brantley, S. T., Smith, C. R., Qi, J., & Hicks, D. W. (2021). Forest restoration increases isolated wetland hydroperiod: a long-term case study. *Ecosphere*, *12*(5), e03495.

Gonzalez A, Lawton JH, Gilbert FS, Blackburn TM, Evans-Freke I (1998) Metapopulation dynamics, abundance, and distribution in a microecosystem. Science, 281, 2045-2047

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., ... & Nusbaum, C. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature biotechnology*, *27*(2), 182-189.

Gompert Z, Forister ML, Fordyce JA, Nice CC, Williamson RJ, Buerkle CA. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of Lycaeides butterflies. *Mol. Ecol.* 19:2455–73 37.

Gorman, T. A., Haas, C. A., & Himes, J. G. (2013). Evaluating methods to restore amphibian habitat in fire-suppressed pine flatwoods wetlands. *Fire Ecology*, *9*(1), 96-109.

Gorman, T. A., Powell, S. D., Jones, K. C., & Haas, C. A. (2014). Microhabitat characteristics of egg deposition sites used by reticulated flatwoods salamanders. *Herpetological Conservation and Biology*, *9*(3), 543-550.

Goslee, S., Urban, D., & Goslee, M. S. (2020). Package 'ecodist'. *Package 'ecodist'*.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular ecology notes*, *5*(1), 184-186.

Graeter, G. J., Rothermel, B. B., & Gibbons, J. W. (2008). Habitat selection and movement of pond-breeding amphibians in experimentally fragmented pine forests. *The Journal of Wildlife Management*, *72*(2), 473-482.

Graham S, Steen D, Nelson K, Durso A, Maerz J (2010) An overlooked hotspot? Rapid biodiversity assessment reveals a region of exceptional herpetofaunal richness in the Southeastern United States. Southeastern Naturalist, 9(1), 19-34.

Grant, R. A., Chadwick, E. A., & Halliday, T. (2009). The lunar cycle: a cue for amphibian reproductive phenology?. *Animal Behaviour*, *78*(2), 349-357.

Grant, E. H. C., Miller, D. A. W., Schmidt, B. R., Adams, M. J., Amburgey, S. M., Chambert, T., Cruickshank, S. S., Fisher, R. N., Green, D. M., Hossack, B. R., Johnson, P. T. J., Joseph, M. B., Rittenhouse, T. A. G., Ryan, M. E., Waddle, J. H., Walls, S. C., Bailey, L. L., Fellers, G. M., Gorman, T. A., ... Muths, E. (2016). Quantitative evidence for the effects of multiple drivers on continental-scale amphibian declines. Scientific Reports, 6(1), 1– 9

Greenwald KR (2010) Genetic data in population viability analysis: case studies with ambystomatid salamanders. Animal Conservation, 13, 115-122

Gregory, T. R. (2018). Animal genome size database. *http://www. genomesize. com*.

Guerry A, Hunter Jr. M (2002) Amphibian distribution in a landscape of forests and agriculture: an examination of landscape composition and configuration. Conservation Biology, 16, 745-754.

Guo, Y., Long, J., He, J., Li, C. I., Cai, Q., Shu, X. O., ... & Li, C. (2012). Exome sequencing generates high quality data in non-target regions. *BMC genomics*, *13*(1), 1-10.

Gutiérrez-Rodríguez, J., Gonçalves, J., Civantos, E., & Martínez-Solano, I. (2017). Comparative landscape genetics of pond-breeding amphibians in Mediterranean temporal wetlands: The positive role of structural heterogeneity in promoting gene flow. *Molecular Ecology*, *26*(20), 5407-5420.

Hale, M. L., Burg, T. M., & Steeves, T. E. (2012). Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies.

Halliday, T. R. (2008). Why amphibians are important. *International Zoo Yearbook*, *42*(1), 7-14.

Hamer A, McDonnell M (2008) Amphibian ecology and conservation in the urbanizing world: A review. Biological Conservation, 141, 2432-2449.

Hanski, I. (1994). A practical model of metapopulation dynamics. *Journal of animal ecology*, 151-162.

Harper, E. B., Rittenhouse, T. A., & Semlitsch, R. D. (2008). Demographic consequences of terrestrial habitat loss for pool-breeding amphibians: predicting extinction risks associated with inadequate size of buffer zones. *Conservation Biology*, *22*(5), 1205-1215.

Harrison, S., Murphy, D. D., & Ehrlich, P. R. (1988). Distribution of the bay checkerspot butterfly, Euphydryas editha bayensis: evidence for a metapopulation model. *The American Naturalist*, *132*(3), 360-382.

Harrison, S. (1991). Local extinction in a metapopulation context: an empirical evaluation. *Biological journal of the Linnean Society*, *42*(1-2), 73-88.

Harrison, S., & Hastings, A. (1996). Genetic and evolutionary consequences of metapopulation structure. *Trends in Ecology & Evolution*, *11*(4), 180-183.

Harrisson, K. A., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. Evolutionary Applications, 7, 1008–1025. https://doi.org/10.1111/eva.12149

Harrisson, K. A., Amish, S. J., Pavlova, A., Narum, S. R., Telonis-Scott, M., Rourke, M. L., ... & Sunnucks, P. (2017). Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. *Molecular ecology*, *26*(22), 6253-6269.

Hartl D, Clark A (1997) Principle of population genetics. Sinauer Associates, Sunderland.

Hedrick, P. W. (2005). A standardized genetic differentiation measure. *Evolution*, *59*(8), 1633-1638.

Hels T, Nachman G (2002) Simulating viability of a spadefoot toad Pelobates fuscus metapopulation in a landscape fragmented by a road. Ecography, 25, 730-744.

Heyer, R., Donnelly, M. A., Foster, M., & Mcdiarmid, R. (Eds.) (1994). Measuring and monitoring biological diversity, standard methods for amphibians. Smithsonian Institution

Hiers, J. K., Laine, S. C., Bachant, J. J., Furman, J. H., Greene Jr, W. W., & Compton, V. (2003). Simple spatial modeling tool for prioritizing prescribed burning activities at the landscape scale. *Conservation Biology*, *17*(6), 1571-1578.

Huff, D.D., Miller, L.M., Chizinski, C.J. & Vondracek, B. (2011). Mixed-source reintroductions lead to outbreeding depression in second-generation descendents of a native North American fish. Mol. Ecol. 20, 4246–4258

Hodges, E., Rooks, M., Xuan, Z., Bhattacharjee, A., Benjamin Gordon, D., Brizuela, L., ... & Hannon, G. J. (2009). Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. *Nature protocols*, *4*(6), 960-974.

Hoffberg, S. L., Kieran, T. J., Catchen, J. M., Devault, A., Faircloth, B. C., Mauricio, R., & Glenn, T. C. (2016). RAD cap: sequence capture of dual-digest RAD seq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, *16*(5), 1264-1278.

Hoffmann AA, Sgró CM. 2011. Climate change and evolutionary adaptation. *Nature*, 470:479–85 49.

Hoffmann AA, Hallas RJ, Dean JA, Schiffer M (2003) Low potential for climatic stress adaptation in a rainforest *Drosophila* species. *Science*, **301**, 100– 102.

Hoffmann, A., Griffin, P., Dillon, S., Catullo, R., Rane, R., Byrne, M., ... & Sgrò, C. (2015). A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses*, *2*(1), 1-24.

Höglund, J., J. K. Larsson, D. Baines, C. Corrales, G. Santafe, and G. Segelbacher 2011. Genetic structure among black grouse in Britain: implications for designing conservation units. *Animal Conservation* **14**: 400– 408.

Houlahan, J., Findlay, C. S., Schmidt, B. R., Meyer, A. H., & Kuzmin, S. L. (2000). Quantitative evidence for global amphibian population declines. *Nature*, *404*(6779), 752-755.

Hutchison, D. W., & Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, *53*(6), 1898-1914. Jacquemyn, H., Brys, R., & Hermy, M. (2002). Patch occupancy, population size and reproductive success of a forest herb (Primula elatior) in a fragmented landscape. *Oecologia*, *130*(4), 617-625.

Inventory, F. N. A. (2022). Southeast Longleaf Pine Ecosystem Occurrences Geodatabase. *Communities*.

Joly P, Miaud C, Lehmann A, Grolet O (2001) Habitat matrix effects on pond occupancy in newts. Conservation Biology, 15, 239-248.

Jones, K. C., Gorman, T. A., Rincon, B. K., Allen, J., Haas, C. A., & Engeman, R. M. (2018). Feral swine Sus scrofa: a new threat to the remaining breeding wetlands of the Vulnerable reticulated flatwoods salamander Ambystoma bishopi. *Oryx*, *52*(4), 669-676.

Jones, O. R., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular ecology resources*, *10*(3), 551-555.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Kapusta, A., Suh, A., & Feschotte, C. (2017). Dynamics of genome size evolution in birds and mammals. *Proceedings of the National Academy of Sciences*, *114*(8), E1460-E1469.

Kardos, M., Luikart, G., & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity*, *115*(1), 63-72.

Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016). Genomics advances the study of inbreeding depression in the wild. *Evolutionary applications*, *9*(10), 1205-1218.

Kelly, E., & Phillips, B. L. (2016). Targeted gene flow for conservation. *Conservation Biology*, **30**, 259– 267. https://doi.org/10.1111/cobi.12623

Kingsolver, J. G., & Pfennig, D. W. (2007). Patterns and power of phenotypic selection in nature. *Bioscience*, *57*(7), 561-572.

Koen, E. L., Bowman, J., & Wilson, P. J. (2015). Node-based measures of connectivity in genetic networks. *Molecular Ecology Resources*, *16*(1), 69–79. https://doi.org/10.1111/1755-0998.12423

Komoroske, L. M., Miller, M. R., O'Rourke, S. M., Stewart, K. R., Jensen, M. P., & Dutton, P. H. (2019). A versatile Rapture (RAD-Capture) platform for genotyping marine turtles. *Molecular ecology resources*, *19*(2), 497-511.

Kreyer D, Oed A, Ealther-Hallwig K, Frank R (2004). Are forests potential landscape barriers for foraging bumblebees? Landscape scale experiments with *Bombus terrestris* agg. and *Bombus pascuorum* (Hymenoptera, Apidae). *Biol Conserv* **116**: 111–118.

Laikre, L., Schwartz, M. K., Waples, R. S., Ryman, N., & GeM Working Group. (2010). Compromising genetic diversity in the wild: unmonitored large-scale release of plants and animals. *Trends in ecology & evolution*, *25*(9), 520-529.

Larsson, J. K., H. A. H. Jansman, G. Segelbacher, J. Höglund, and H. P. Koelewijn 2008. Genetic impoverishment of the last black grouse (*Tetrao tetrix*) population in the Netherlands: detectable only with a reference from the past. *Molecular Ecology* **17**: 1897– 1904.

Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., ... & Buckler, E. S. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Science advances*, *1*(6), e1400218.

Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., ... & Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness—Implications for brown trout conservation. *Ecology and Evolution*, *9*(4), 2106-2120.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, *25*(14), 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Loredo, I., Van Vuren, D., & Morrison, M. L. (1996). Habitat use and migration behavior of the California tiger salamander. *Journal of Herpetology*, *30*(2), 282-285.

Luikart, G., Kardos, M., Hand, B. K., Rajora, O. P., Aitken, S. N., & Hohenlohe, P. A. (2018). Population genomics: advancing understanding of nature. In *Population genomics* (pp. 3-79). Springer, Cham.

Luikart, G., Antao, T., Hand, B. K., Muhlfeld, C. C., Boyer, M. C., Cosart, T., ... & Waples, R. S. (2021). Detecting population declines via monitoring the effective number of breeders (Nb). *Molecular Ecology Resources*, *21*(2), 379-393.

Luu, K., Bazin, E., & Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular ecology resources*, *17*(1), 67-77.

Madison DM (1997) The emigration of radio-implanted spotted salamanders, Ambystoma maculatum. Journal of Herpetology, 31, 542-551.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in ecology & evolution*, *18*(4), 189-197.

Manel S, Joost S, Epperson BK, Holderegger R, Storfer A, et al. 2010. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Mol. Ecol.* 19:3760–72 47.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614–621. https://doi.org/10.1016/j.tree.2013.05.012

Marsh DM, Trenham PC (2008) Metapopulation dynamics and amphibian conservation. *Conservation Biology*, **15**, 40–49.

Martin, K. L. (1999). Ready and waiting: delayed hatching and extended incubation of anamniotic vertebrate terrestrial eggs. *American Zoologist*, *39*(2), 279-288.

Martinez-Solano, I., & Gonzalez, E. G. (2008). Patterns of gene flow and source-sink dynamics in high altitude populations of the common toad Bufo bufo (Anura: Bufonidae). *Biological Journal of the Linnean Society*, *95*(4), 824-839.

Mayhew, W. W. (1965). Adaptations of the amphibian, Scaphiopus couchi, to desert conditions. *American Midland Naturalist*, 95-109.

McCallum, M. L. (2007). Amphibian decline or extinction? Current declines dwarf background extinction rate. *Journal of Herpetology*, *41*(3), 483-491.

McCartney-Melstad, E., Vu, J. K., & Shaffer, H. B. (2018). Genomic data recover previously undetectable fragmentation effects in an endangered amphibian. Molecular Ecology, 27(22), 4430– 4443.

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, **17**, 285– 291.

McMahon, B. J., M. P. Johansson, S. B. Piertney, K. Buckley, and J. Höglund 2012. Genetic variation among endangered Irish red grouse (*Lagopus lagopus hibernicus*) populations: implications for conservation and management. *Conservation Genetics* **13**: 639– 647.

Means, R. P. (2013). *Ambystoma bishopi: Survey and Status of Designated Critical Habitat.* Coastal Plains Institute. https://www.coastalplains.org/wp-content/uploads/2018/10/Ambystoma-bishopi-Survey-and-Status-of-Designated-Critical-Habitat.pdf

Meyer, J. L., Strayer, D. L., Wallace, J. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks 1. *JAWRA Journal of the American Water Resources Association*, *43*(1), 86-103.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome research*, *17*(2), 240-248.

Mitchell J, Gibbons W (2010) Salamanders of the Southeast. University of Georgia Press, Athens.

Moritz, C. (1999). Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas*, *130*(3), 217-228.

Murphy, M. A., Dezzani, R., Pilliod, D. S., & Storfer, A. (2010). Landscape genetics of high mountain frog metapopulations. *Molecular ecology*, *19*(17), 3634-3649.

Murphy, M. A., & Evans, J. S. (2011). Genetic patterns as a function of landscape process: Applications of neutral genetic markers for predictive modeling in landscape ecology. In *Predictive species and habitat modeling in Landscape Ecology* (pp. 161-188). Springer, New York, NY.

Murphy, M. O., Jones, K. S., Price, S. J., & Weisrock, D. W. (2018). A genomic assessment of population structure and gene flow in an aquatic salamander identifies the roles of spatial scale, barriers, and river architecture. *Freshwater Biology*, *63*(5), 407-419.

Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources*, *17*(6), 1136-1147.

Nowoshilow, S., Schloissnig, S., Fei, JF. *et al.* The axolotl genome and the evolution of key tissue formation regulators. *Nature* **554,** 50–55 (2018). https://doi.org/10.1038/nature25458

Nunziata, S. O., & Weisrock, D. W. (2018). Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, *120*(3), 196-207.

O'Donnell KM, Messerman AF, Barichivich WJ et al (2017) Structured decision making as a conservation tool for recovery planning of two endangered salamanders. J Nat Conserv 37:66–72.

Oksanen, J., F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner (2019). vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists.

Palis, J. G. (1995). Larval growth, development, and metamorphosis of Ambystoma cingulatum on the Gulf Coastal Plain of Florida. *Florida Scientist*, 352-358.

Palis J (1997) Distribution, habitat, and status of the Flatwoods Salamander (*Ambystoma cingulatum*) in Florida, USA. Herpetological Natural History 5:53-65

Palis, J. G. (1997). Breeding migration of Ambystoma cingulatum in Florida. *Journal of Herpetology*, 71-78.

Palis JG, Aresco MJ, Kilpatrick S (2006) Breeding biology of a Florida population of *Ambystoma cingulatum* (Flatwoods salamander) during a drought. Southeastern Naturalist 5(1):1-8

Palis, J. G., Consul, P. E., & Enge, K. (2006). Capture Rate, Body Size, and Survey Recommendations for Larval Ambystoma cingulatum (Flatwoods... *SOUTHEASTERN NATURALIST*, *5*(1), 9-16.

Palis J, Hammerson G (2008) Ambystoma bishopi. The IUCN Red List of Threatened Species 2008. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T136128A4245386.en.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526-528.

Pauly G, Piskurek O, Shafer H (2007) Phylogeographic concordance in the southeastern United States: the fatwoods salamander, Ambystoma cingulatum, as a test case. Mol Ecol 16:415–429

Pauly, G. B., Bennett, S. H., Palis, J. G., & Bradley Shaffer, H. (2012). Conservation and genetics of the frosted flatwoods salamander (Ambystoma cingulatum) on the Atlantic coastal plain. *Conservation Genetics*, *13*(1), 1-7.

Pearman PB, Garner TWJ (2006) Susceptibility of Italian agile frog populations to an emerging strain of Ranavirus parallels population genetic diversity. Ecology Letters, 8, 401-408.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559-572.

Peterman, W. E., Crawford, J. A., & Semlitsch, R. D. (2008). Productivity and significance of headwater streams: population structure and biomass of the black-bellied salamander (Desmognathus quadramaculatus). *Freshwater Biology*, *53*(2), 347-357.

Peterman, W. E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, *9*(6), 1638-1647.

Petranka JW, Holbrook CT (2006) Wetland restoration for amphibians: should local sites be designed to support metapopulations or patchy populations? *Restoration Ecology*, **14**, 404– 411.

Pfeiffer, F., Gröber, C., Blank, M., Händler, K., Beyer, M., Schultze, J. L., & Mayer, G. (2018). Systematic evaluation of error rates and causes in short samples in next-generation sequencing. Scientific Reports, 8(1), 10950. https://doi.org/10.1038/s41598-018-29325-6

Pflüger, F. J., & Balkenhol, N. (2014). A plea for simultaneously considering matrix quality and local environmental conditions when analysing landscape impacts on effective dispersal. *Molecular Ecology*, *23*(9), 2146–2156. https://doi.org/10.1111/mec.12712

Pierson, J. C., Coates, D. J., Oostermeijer, J. G. B., Beissinger, S. R., Bragg, J. G., Sunnucks, P., ... & Young, A. G. (2016). Genetic factors in threatened species recovery plans on three continents. *Frontiers in Ecology and the Environment*, *14*(8), 433-440.

Pilliod, D. S., Peterson, C. R., & Ritson, P. I. (2002). Seasonal migration of Columbia spotted frogs (Rana luteiventris) among complementary resources in a high mountain basin. *Canadian Journal of Zoology*, *80*(11), 1849-1862.

Pilot, M., Greco, C., vonHoldt, B. *et al.* Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. *Heredity* **112,** 428–442 (2014). https://doi.org/10.1038/hdy.2013.122

Pisa, G., Orioli, V., Spilotros, G., Fabbri, E., Randi, E., & Bani, L. (2015). Detecting a hierarchical genetic population structure: the case study of the Fire Salamander (Salamandra salamandra) in Northern Italy. *Ecology and Evolution*, *5*(3), 743-758.

Powell, S. D., Gorman, T. A., & Haas, C. A. (2015). Nightly movements and use of burrows by reticulated flatwoods salamanders (Ambystoma bishopi) in breeding wetlands. *Florida Scientist*, 149-155.

Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2015). Fishing for selection, but only catching bias: 1.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Pritchard, J. K., Wen, X., & Falush, D. (2010). Documentation for STRUCTURE software, version 2.3. University of Chicago, Chicago, IL.

Pruett, C., & Winker, K. (2008). The effects of sample size on population genetic diversity estimates in song sparrows Melospiza melodia. *Journal of Avian Biology*, *39*(2), 252-256.

Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular ecology resources*, *16*(3), 608-627.

Pulliam, H. R. (1988). Sources, sinks, and population regulation. *The American Naturalist*, *132*(5), 652-661.

Purrenhage, J. L., Niewiarowski, P. H., & Moore, F. G. (2009). Population structure of spotted salamanders (Ambystoma maculatum) in a fragmented landscape. *Molecular Ecology*, *18*(2), 235-247.

Quinn H. (1980) Captive propagation of endangered Houston toads. *Herpetological Review*, 11, 109.

Ranvestel, A. W., Lips, K. R., Pringle, C. M., Whiles, M. R., & Bixby, R. J. (2004). Neotropical tadpoles influence stream benthos: evidence for the ecological consequences of decline in amphibian populations. *Freshwater Biology*, *49*(3), 274-285.

Reed, D. H., & Frankham, R. (2003). Correlation between Fitness and Genetic Diversity. *Conservation Biology*, *17*(1), 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348-4370.

Richardson, J. L., & Urban, M. C. (2013). Strong selection barriers explain microgeographic adaptation in wild salamander populations. *Evolution*, *67*(6), 1729-1740.

Riley S, Busteed GT, Kats LB, Vandergon TL, Lee LFS, Dagit RG, Kerby JL, Fisher RN, Sauvajot RM (2005) Effects of urbanization on the distribution and abundance of amphibians and invasive species in Southern California streams. Conservation Biology, 19(6), 1894-1907.

Robinson JD, Moyer GR (2013) Linkage disequilibrium and effective population size when generations overlap. Evolutionary Applications, 6(2), 290-302.

Rowe, G., & Beebee, T. J. C. (2004). Reconciling genetic and demographic estimators of effective population size in the anuran amphibian Bufo calamita. *Conservation Genetics*, *5*(3), 287-298.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Ryman, N., & Laikre, L. (1991). Effects of supportive breeding on the genetically effective population size. *Conservation Biology*, *5*(3), 325-329.

Sagvik, J., Uller, T. & Olsson, M. Outbreeding depression in the common frog, *Rana temporaria*. *Conserv Genet* **6,** 205–211 (2005). https://doi.org/10.1007/s10592-004-7829-3

Salice C (2012) Multiple stressors and amphibians: contributions of adverse health effects and altered hydroperiod to population decline and extinction. Journal of Herpetology, 46(4), 675-681.

Savage, A. E., Muletz-Wolz, C. R., Campbell Grant, E. H., Fleischer, R. C., & Mulder, K. P. (2019). Functional variation at an expressed MHC class IIβ locus associates with Ranavirus infection intensity in larval anuran populations. *Immunogenetics*, *71*(4), 335-346.

Schindler, D. E., Hilborn, R., Chasco, B., Boatright, C. P., Quinn, T. P., Rogers, L. A., & Webster, M. S. (2010). Population diversity and the portfolio effect in an exploited species. *Nature*, *465*(7298), 609-612.

Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in ecology & evolution*, *22*(1), 25-33.

Semlitsch, R. D. (1987). Interactions between fish and salamander larvae. *Oecologia*, *72*(4), 481-486.

Semlitsch, R. D., & Anderson, T. L. (2016). Structure and dynamics of spotted salamander (Ambystoma maculatum) populations in Missouri. *Herpetologica*, *72*(2), 81-89.

Semlitsch RD (2008) Differentiating migration and dispersal processes for pond-breeding amphibians. *Journal of Wildlife Management*, **72**, 260– 267.

Semlitsch, R. D., Scott, D. E., & Pechmann, J. H. (1998). Time and size at metamorphosis related to adult fitness in Ambystoma talpoideum. *Ecology*, *69*(1), 184-192.

Semlitsch, R. D., D. E. Scott, J. H. K. Pechmann, and J. W. Gibbons. 1996. Structure and dynamics of an amphibian community: evi- dence from a 16-year study of a natural pond. Pages 217-247 in M. L. Cody and J. A. Smallwood, editors. Long-term studies of verte- brate communities. Academic Press, San Diego, California.

Semlitsch, R. D., Walls, S. C., Barichivich, W. J., & O'Donnell, K. M. (2017). Extinction debt as a driver of amphibian declines: an example with imperiled flatwoods salamanders. *Journal of Herpetology*, *51*(1), 12-18.

Sessions SK. Evolutionary cytogenetics in salamanders. Chromosom Res. 2008;16:183–201.

Shafer, A. B., Wolf, J. B., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., ... & Fawcett, K. D. (2015). Genomics and the challenging translation into conservation practice. *Trends in ecology & evolution*, *30*(2), 78-87.

Skelly, D. K., Werner, E. E., & Cortwright, S. A. (1999). Long-term distributional dynamics of a Michigan amphibian assemblage. *Ecology*, *80*(7), 2326-2337.

Skriver, P. (1988). A pond restoration project and a tree-frog(Hyla arborea) project in the municipality of Aarhus. *Memoranda Societatis pro Fauna et Flora Fennica*, *64*(3), 146-147.

Slate, J., David, P., Dodds, K. G., Veenvliet, B. A., Glass, B. C., Broad, T. E., & McEwan, J. C. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity*, *93*(3), 255-265.

Smith, M., & M. Green, D. (2005). Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations?. *Ecography*, *28*(1), 110-128.

Smith, J. J., Putta, S., Zhu, W., Pao, G. M., Verma, I. M., Hunter, T., ... & Voss, S. R. (2009). Genic regions of a large salamander genome contain long introns and novel genes. *BMC genomics*, *10*(1), 1-11.

Spear SF, Peterson CS, Matocq MD, Storfer A (2005) Landscape genetics of the blotched tiger salamander (Ambystoma tigrinum melanostictum). Molecular Ecology, 14, 2553-2564.

Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, **18**, 94– 101.

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., ... & Waits, L. P. (2007). Putting the 'landscape'in landscape genetics. *Heredity*, *98*(3), 128-142.

Strand, T., G. Segelbacher, M. Quintela, L. Xiao, T. Axelsson, and J. Höglund 2012. Can balancing selection on MHC loci counteract genetic drift in small fragmented populations of black grouse? *Ecology and Evolution* **2**: 341– 353.

Stuart S, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, Waller RW (2004) Status and trends of amphibian declines and extinctions worldwide. Science, 306, 1783- 1786.

Sun, C., Shepard, D. B., Chong, R. A., López Arriaza, J., Hall, K., Castoe, T. A., ... & Mueller, R. L. (2012). LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. *Genome biology and evolution*, *4*(2), 168-183.

Sun, C., & Mueller, R. L. (2014). Hellbender genome sequences shed light on genomic expansion at the base of crown salamanders. *Genome biology and evolution*, *6*(7), 1818-1829.

Sunny, A., Monroy-Vilchis, O., Reyna-Valencia, C., & Zarco-González, M. M. (2014). Microhabitat types promote the genetic structure of a micro-endemic and critically endangered mole salamander (Ambystoma leorae) of Central Mexico. *PLoS One*, *9*(7), e103595.

Taylor, B. E., Scott, D. E., & Gibbons, J. W. (2006). Catastrophic reproductive failure, terrestrial survival, and persistence of the marbled salamander. *Conservation Biology*, *20*(3), 792-801.

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*, *14*(2), 178-192.

USFWS [U.S. Fish and Wildlife Service] (2015) Reticulated flatwoods salamander (Ambystoma bishopi) 5-year review: summary and evaluation. Federal Register. 79 FR 56821

USFWS [U.S. Fish and Wildlife Service] (2020) Species Status Assessment for the Reticulated Flatwoods Salamander (*Ambystoma bishopi*).

USFWS [U.S. Fish and Wildlife Service] (2021) Recovery Plan For the Reticulated Flatwoods Salamander (*Ambystoma bishopi*).

Van Lear, D. H., Carroll, W. D., Kapeluck, P. R., & Johnson, R. (2005). History and restoration of the longleaf pine-grassland ecosystem: implications for species at risk. *Forest ecology and Management*, *211*(1-2), 150-165.

Vandersteen Tymchuk W, O'Reilly P, Bittman J, MacDonald D, Schulte P. 2010. Conservation genomics of Atlantic salmon: variation in gene expression between and within regions of the Bay of Fundy. *Mol. Ecol.* 19:1842–59 48.

vonHoldt BM, Pollinger JP, Earl DA, Knowles JC, Boyko AR, et al. 2011. A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Res*. 21:1294–305

Walker RS, Novaro AJ, Branch LC (2003). Effects of patch attributes, barriers, and distance between patches on the distribution of a rock-dwelling rodent (*Lagidium viscacia*). *Landscape Ecol* **18**: 187–194.

Walls SC, Waddle JH, Faulkner SP (2014) Wetland reserve program enhances site occupancy and species richness in assemblages of anuran amphibians in the Mississippi Alluvial Valley, USA. Wetlands, 34(1), 197-207.

Wagner, H.H., Fortin, MJ. A conceptual framework for the spatial analysis of landscape genetic data. *Conserv Genet* **14,** 253–261 (2013). https://doi.org/10.1007/s10592-012-0391-5

Wang, I. J., Savage, W. K., & Bradley Shaffer, H. (2009). Landscape genetics and least-cost path analysis reveal unexpected dispersal routes in the California tiger salamander (Ambystoma californiense). *Molecular ecology*, *18*(7), 1365-1374.

Wang, J., & Ryman, N. (2001). Genetic effects of multiple generations of supportive breeding. *Conservation Biology*, *15*(6), 1619-1631.

Wang J (2016) Individual identification from genetic marker data: developments and accuracy comparisons of methods. Molecular Ecology Resources, 16, 163-175.

Wang IJ, Johnson JR, Johnson BB, Shaffer HB (2011) Effective population size is strongly correlated with breeding pond size in the endangered California tiger salamander, Ambysotma californiense. Conservation Genetics, 12(4), 911-920.

Wang, I. J., & Shaffer, H. B. (2017). Population genetic and field-ecological analyses return similar estimates of dispersal over space and time in an endangered amphibian. *Evolutionary Applications*, *10*(6), 630-639.

Waples RS (2002) Definition and estimation of effective population size in the conservation of endangered species. In: Beissinger SR, McCullough DR (eds) Population viability analysis. University of Chicago Press, Chicago, IL.

Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, *197*(2), 769-780.

Waples, R. S. (2016). Making sense of genetic estimates of effective population size. *Molecular Ecology.*

Watts, A. G., Schlichting, P., Billerman, S., Jesmer, B., Micheletti, S., Fortin, M. J., ... & Murphy, M. A. (2015). How spatio-temporal habitat connectivity affects amphibian genetic structure. *Frontiers in Genetics*, *6*, 275.

Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., ... & Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications*, *4*(6), 709-725.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *evolution*, 1358-1370.

Weisrock, D. W., Hime, P. M., Nunziata, S. O., Jones, K. S., Murphy, M. O., Hotaling, S., & Kratovil, J. D. (2018). Surmounting the large-genome "problem" for genomic data generation in salamanders. In *Population genomics: wildlife* (pp. 115-142). Springer, Cham.

Wells, K. D. (2010). The ecology and behavior of amphibians. University of Chicago Press, Chicago, US.

Wells, K. D. 1977. The social behaviour of anuran amphibians. Animal Behaviour 25:666-693.

Wendt, A., Haas, C. A., Gorman, T., & Roberts, J. H. (2021). Metapopulation genetics of endangered reticulated flatwoods salamanders (Ambystoma bishopi) in a dynamic and fragmented landscape. *Conservation Genetics*, *22*(4), 551-567.

Werner E. E., Gilliam J. F. 1984. The ontogenetic niche and species interactions in size-structured populations. Annual Review of Ecology and Systematics 15:393– 425.

Whiteley, A. R., Coombs, J. A., Cembrola, M., O'Donnell, M. J., Hudy, M., Nislow, K. H., & Letcher, B. H. (2015). Effective number of breeders provides a link between interannual variation in stream flow and individual reproductive contribution in a stream salmonid. *Molecular Ecology*, *24*(14), 3585-3602.

Whitlock, M. C. (1992). Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution*, *46*(3), 608-615.

Whitlock MC, Barton NH (1997) The effective size of a subdivided population. Genetics, 146(1), 427-441.

Wilbur, H. M., & Collins, J. P. (1973). Ecological aspects of amphibian metamorphosis: nonnormal distributions of competitive ability reflect selection for facultative metamorphosis. *Science*, *182*(4119), 1305-1314.

Wilbur H. M. 1980. Complex life cycles. Annual Review of Ecology and Systematics 11:67–93.

Williams, S. T., Elbers, J. P., & Taylor, S. S. (2021). Population structure, gene flow, and sex-biased dispersal in the reticulated flatwoods salamander (Ambystoma bishopi): Implications for translocations. *Evolutionary applications*, *14*(9), 2231-2243.

Willson, J. D., & Dorcas, M. E. (2003). Quantitative sampling of stream salamanders: comparison of dipnetting and funnel trapping techniques. *Herpetological Review*, *34*(2), 128-129.

Wolf, S., Hartl, B., Carroll, C., Neel, M. C., & Greenwald, D. N. (2015). Beyond PVA: Why recovery under the Endangered Species Act is more than population viability. *BioScience*, *65*(2), 200-207.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*(2), 97.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. Trends in Ecology and Evolution, 11(10), 413-418

Zamudio KR, Wieczorek AM (2007) Fine-scale spatial genetic structure and dispersal among spotted salamander (Ambystoma maculatum) breeding populations. Molecular Ecology, 16, 257-274.

Zeisset, I., & Beebee, T. J. C. (2013). Donor population size rather than local adaptation can be a key determinant of amphibian translocation success. *Animal Conservation*, *16*(3), 359-366.

Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape ecology*, *27*(6), 777-797.

APPENDIX

STACKS DEFAULT SETTING APPLIED IN BIOINFORMATIC PIPELINE.

Stacks step	Filter	Default setting	Description
ustacks	max_locus_stacks	3	maximum number of stacks at a single de novo locus
	max_gaps	2	number of gaps allowed between stacks before merging
	alpha	0.05	chi square significance level required to call heterozygote or homozygote
	bound_low	0	lower bound for epsilon, the error rate
	bound_high	1	upper bound for epsilon, the error rate
gstacks	model	marukilow	model to use to call variants and genotypes
	var-alpha	0.05	alpha threshold for discovering SNPs
	gt-alpha	0.05	alpha threshold for calling genotypes