

INFERRING THE EVOLUTIONARY TRAJECTORIES OF WILD, FRAGMENTED  
POPULATIONS USING GENOMIC AND LONG-TERM DEMOGRAPHIC DATA

By

Meaghan Ianna Clark

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

Integrative Biology – Doctor of Philosophy  
Ecology, Evolutionary Biology, and Behavior – Dual Major

2024

## ABSTRACT

Anthropogenic land conversion has dramatically altered the environment experienced by wild populations, often resulting in small and fragmented populations. It is urgent that we develop approaches to understand the ecological and evolutionary dynamics of such populations to craft effective conservation management. In this dissertation, I use genomics, life history information, and long-term demographic data to explore the evolutionary dynamics of small populations. Using simulations, I test how, and under what conditions, comparisons between the population genetics of different age groups can be used to identify population demographic declines (Chapter 2). I discover that older individuals retain higher genetic diversity representative of past population size after a decline, and that comparison to the genetic diversity in younger individuals can increase power to detect a decline when generation time is long, and a decline has been severe. With a focus on the eastern massasauga (*Sistrurus catenatus*), I next explore how anthropogenic land conversion has impacted already small and fragmented populations of this threatened rattlesnake. First, I document population structure and demographic history across the range of eastern massasaugas and test if anthropogenic disturbance is a predictor of inbreeding and genetic drift, with particular attention to populations at their range center (Chapter 3). I find significant structure among populations that follows a pattern of isolation by distance, as well as evidence for recent inbreeding that is associated with increased history of disturbance. Second, using a rich dataset of over ten years of capture-recapture data, I test if inbreeding in eastern massasaugas impacts two major components of fitness: survival and reproduction (Chapter 4). I find that snakes with increased inbreeding have lower apparent annual survival and lower probability of having offspring, direct evidence of inbreeding depression. Together, this body of research highlights the power of genomic and life

history data to understand the hidden dynamics of small and isolated populations.

To Helen L. and David C. Clark, who have always believed in my dreams and ambitions.

## ACKNOWLEDGMENTS

I am incredibly grateful to have been surrounded by strong and collaborative communities throughout my scientific career. So much of this work has been made possible by their intellectual and emotional support, and energized by their curiosity, enthusiasm, humor, and joy. In many ways, my PhD has been defined in pairs: from co-advisement, being part of two lab groups, and learning from two academic communities. As such, this section might be rather long, a reflection of how doubly lucky I have been.

First, I could not have chosen more wonderful and dedicated PhD advisors than Gideon Bradburd and Sarah Fitzpatrick. Gideon, thank you for the time and mentorship you gave a random master's student who wanted to run one of your methods on her frog data. The generosity and kindness you showed me then have continued to be a hallmark of our working relationship. Having had the opportunity to observe and learn from how you approach biological and statistical questions will forever better my research. Sarah, thank you for your constant encouragement and for always having my back, for treating me like a respected colleague as well as your advisee, and for always leading by example. I have learned so much from you both about how to creatively and robustly analyze data, as well as how to navigate the world with care and kindness. Thank you for pushing me to grow and to have confidence in myself and my abilities.

My committee members, Jeff Conner and Fred Janzen, have also played integral parts in my development as a researcher. Jeff has taught me to be a careful and thorough thinker, and to always root my work in the fundamental principles of evolutionary genetics. Fred has always helped me place my research within a broader ecological, historical, and herpetological context. I am a stronger researcher because of their questions, insights, and advice. I am also forever grateful to my eastern massasauga collaborators: Jen Moore, Eric Hileman, Lisa Faust, and

Danielle Bradke, for trusting me with years of hard-earned data and precious samples, and for sharing their knowledge and love of eastern massasaugas. It has been a privilege to be part of this collaboration.

Bradburd and Fitzpatrick labs, where do I even begin? The postdocs and other members of the Bradburd lab have been stellar role models and friends. From backyard bonfires to letting me ask them a million questions, thank you to Matteo Tomasini, Bob Week, Zach Hancock, Leonard Jones, and Mike Grundler, for their wisdom, code troubleshooting, and friendship. I will miss running up hills in the Arboretum with Nicole Adams, not to mention her bioinformatics expertise. Thank you to Rachel Toczydlowski for teaching me to enjoy the sun while it's shining (even if it's shining in my eyes), and a million other things about genomics and friendship. I am also grateful to the EEB community at the University of Michigan, particularly Kristen Wacker, Teresa Pegan, and Abby Kimmitt, for welcoming an occasional visitor into their midst.

It has been a pleasure to watch Alex Lewanski craft such an impressive research portfolio during our time as the bridge between the Bradburd and Fitzpatrick labs. I am incredibly honored and flattered by his occasional description of himself as my 'mini-me'. Kyle Jaynes and Isabela Borges taught me everything I needed to know about graduate student life at KBS and MSU. Thank you for encouraging me to strive for more work-life balance than I ever thought possible and for putting up with my enthusiasm for the diversity of pond lab leeches. Thank you to Fitz lab postdocs Brendan Reid, Cinnamon Mittan-Moreau, Jessica Judson, and Alexa Guerrero for their generosity of time and knowledge and camaraderie in the face of tight deadlines. Tyler Linderoth has been an incredible mentor and true friend. Thank you for spending so many hours with me delving into everything from the assumptions of Watterson's estimator of theta, introducing me to the world of professional bass fishing, and inspiring a musical revolution

within the lab.

At MSU, I have felt welcomed by the KBS, IBIO, and EEB communities. Thank you in particular to Marjorie Weber, Emily Josephs, Acer VanWallendael, and Joe Riedy. I am grateful for the many friends I have made in the broader MSU community: Olivia Fitch, Emily Dean, Erika LaPlante, Connie Rojas, Olivia Utley, Moriah Young, Bruce Martin, Eleanore Ritter, Maya Wilson Brown, Brooke Jeffrey, Begoña Berenguer, Jamie Smith, Sylvie Martin-Eberhardt, and many others. Thank you for the many brunches, pandemic zoom workouts, movie nights, and Dungeons and Dragons campaigns.

Many of the professional connections I have today are forged by a shared enthusiasm for working with Jeanne Robertson. I so appreciate her continued mentorship, wisdom, and patience. I am grateful for the continued support of mentors and friends from California State University, Northridge: Robert Espinoza, Andres Vega, Maria Akopyan, Emma Collosi, Sarah Wenner, Carson Keller, Andrew Powers, and Rachel Larson. Previous mentors from my undergraduate and lab technician days helped me prepare for graduate school and my career beyond: Lisette Waits, Paul Hohenlohe, Kim Andrews, Jen Adams, Kathleen O'Malley, Ben Duffus, Matt Crook, Delbert Hutchison, Heidi Dobson, and Arielle Cooley. Thank you to Sarah Hendricks and Amanda Stahlke for showing me that being a graduate student means working hard, generating exciting ideas, and being an active participant in your community. Whether by a twist of fate or attending a college of majority biology majors, I have been lucky to journey through grad school alongside Katie Jenike and Mara Heilig. Thank you for all your advice and mountain adventures.

I would like to thank my family for their love and their evergreen support of my education. I am grateful to my grandmother Julie Lindstrom, as well as Lowell and Jen Lindstrom and their family, for providing me a home-away-from-home in the Midwest. Thank

you to Lindy Lindstrom for translating ‘academia’ for me throughout my education.

My grandparents, John David and Margaret Clark helped support my undergraduate degree. I share my appreciation of education with my grandfather, Bill Lindstrom, and my love of the natural world with my uncle, David Lindstrom. I know they would all be proud of me for this accomplishment.

Thank you to my brother, David Maxwell Clark for always making me laugh and helping me debug code (though rarely at the same time). My parents, Helen and David Clark told me that I could do anything I set my mind to, and for better or worse, I believed them. They have provided me with unconditional love and support throughout, and I will be forever grateful and proud to be their daughter.

## TABLE OF CONTENTS

CHAPTER ONE: Introduction .....	1
CHAPTER TWO: Pitfalls and windfalls of detecting demographic declines using population genetics in long-lived species .....	6
CHAPTER THREE: Anthropogenic land use changes associated with increased inbreeding in a threatened rattlesnake.....	30
CHAPTER FOUR: Inbreeding reduces fitness in spatially structured populations of a threatened rattlesnake .....	52
REFERENCES .....	74
APPENDIX A: Supporting information: Pitfalls and windfalls of detecting demographic declines using population genetics in long-lived species.....	90
APPENDIX B: Supporting information: Anthropogenic land use changes associated with increased inbreeding in a threatened rattlesnake .....	92
APPENDIX C: Supporting information: Inbreeding reduces fitness in spatially structured populations of a threatened rattlesnake.....	98

## CHAPTER ONE:

### Introduction

Wild populations face many current threats, including climate change, disease, and habitat loss and fragmentation (Haddad et al., 2015; McCarty, 2001; Pimm et al., 2014). Changing climatic conditions, propelled by anthropogenic CO<sub>2</sub> emissions, are driving vast range shifts (Pham et al., 2022; Ralston & Kirchman, 2013), adaptation (Bi et al., 2019; Hoffmann & Sgrò, 2011), and decline (Bay et al., 2018; R. W. Clark et al., 2011). Concurrently, much of the world's land area has been converted to human land use types, like crop land, pastures, and urban areas (X. Li et al., 2023; Sohl et al., 2016). While these areas can provide suitable habitat for some species, others cannot thrive in anthropogenic land use types. The loss and increasing isolation of habitat, exacerbated by climate change, leads to population decline and loss of neutral and adaptive genetic diversity (Gauthier et al., 2020; Yineger et al., 2014).

Small populations isolated by habitat loss and fragmentation are inherently more susceptible to small population problems such as Allee effects, loss of genetic diversity due to increased rates of genetic drift, and increased inbreeding and risk of inbreeding depression. Allee effects, processes caused by low population density that lead to reduced fitness, are found across taxonomic groups, and can be generated by mate limitation, cooperative defense, predator satiation, cooperative feeding, dispersal, and habitat alternation (Kramer et al., 2009). Genetic drift, the stochastic change in population allele frequencies from one generation to the next, is stronger in small compared to large populations. This is especially consequential when adaptive, or potentially adaptive alleles are lost due to drift, and when deleterious alleles become fixed (Segelbacher et al., 2014). Another problem small populations face is increased inbreeding, or close-kin mating, and the concomitant risk of inbreeding depression. Inbreeding depression

refers to a negative relationship between inbreeding and fitness; it is generated by the expression of deleterious recessive alleles and loss of heterozygosity for overdominant alleles (D. Charlesworth & Willis, 2009; Keller & Waller, 2002). One well-known case study of inbreeding in a small population is that of wolves on Isle Royale in Lake Superior, Michigan, USA. Wolves on Isle Royal in the late 1990s showed a high degree of inbreeding, and signs of inbreeding depression due to increased homozygosity and large-effect deleterious recessive alleles (Robinson et al., 2019). Small population problems can cause populations to enter an “extinction vortex,” where the combined effects of inbreeding, drift, and Allee effects cause populations to become smaller, which further exacerbates small population problems (Gilpin & Soulé, 1986). Without management intervention, this process eventually leads to population extinction. For example, inbreeding depression eventually led to wolves being functionally extirpated from Isle Royale until additional managed translocations (Hoy, Hedrick, et al., 2023).

Genomic sequencing allows researchers to glimpse the history and current genetic status of populations. Patterns within the genome can provide insights into long-term and contemporary effective population size ( $N_e$ ), which reflects the strength of genetic drift in a population (R. S. Waples, 2016, 2022). Genomic data can also be used to measure genetic differentiation between populations (Weir & Cockerham, 1984), which can reflect geographic structure, barriers to dispersal, and demographic history and inform conservation management. Genomic data can also be used to measure inbreeding within populations. By linking genomic data with long-term monitoring data, we can learn about inbreeding depression and the impact of life history on the genome.

In my dissertation, I use simulated and empirical genomic data to learn about small population problems through the lenses of simulations and eastern massasauga rattlesnakes

(*Sistrurus catenatus*). Simulations inform our understanding of evolutionary processes because they allow for the testing of theory and methods in situations where the true underlying dynamics are known. Recent advances in simulation software have allowed for simulations crafted to mimic key life history traits of empirical populations, like generation time, spatial structure, and interactions between species (Haller & Messer, 2019, 2023). Simulations have great power to explore new methods and inform our expectations for empirical studies.

In **Chapter 2** of my dissertation, I use evolutionary simulations to explore the relationship between age and genetic diversity after demographic declines in species with long lifespans and overlapping generations. Specifically, I test if comparing older individuals in a population to younger individuals can be used as pseudo-temporal sampling when temporal sampling is not available. I find there are patterns of genetic diversity indicative of population decline: namely, that older individuals have higher genomic diversity, reflecting higher  $N_e$  at the time of their birth, compared to younger individuals born after a decline in  $N_e$ . Comparing genetic diversity between older and younger individuals after a demographic decline can have equal or greater power compared to temporal sampling when generation time is long. However, these comparisons work best when the demographic decline has been severe. These results give context to empirical studies that have found mixed support for a relationship between age and genetic diversity after declines (Schmidt et al., 2018; Vranckx et al., 2012).

In Chapters 3 and 4 of my dissertation, I leverage genomic sequencing to learn about small population problems in eastern massasaugas. Eastern massasaugas are a threatened species with specific seasonal habitat requirements: they summer in open wetlands and return to upland forests in the fall to overwinter. Key threats to eastern massasauga populations include habitat succession and habitat loss (Moore & Gillingham, 2006; Szymanski et al., 2016). Eastern

massasaugas have long been theorized to exist in small, isolated populations due to their limited dispersal distance and the naturally patchy distribution of their habitat. Previous genomic research found that populations of eastern massasaugas at their range edges in Ohio have existed at small population sizes historically but have also experienced recent declines in population size over the past 200 years, the approximate timing of widescale anthropogenic land cover conversion (Sovic et al., 2019). While range-wide declines and population extirpations have been documented, the range center in Michigan (USA) is seemingly a stronghold for eastern massasaugas (Faust et al., 2011; Szymanski et al., 2016). However, the demographic history and genetic health of Michigan eastern massasaugas is largely unknown.

In **Chapter 3**, I use whole genome sequencing from across the range of eastern massasaugas to document patterns of genetic diversity and inbreeding, and test if populations that have experienced greater habitat disturbance experience increased inbreeding and genetic drift. I find strong population structure across the range of eastern massasaugas in accordance with a pattern of isolation by distance. Populations in northern Michigan are most differentiated, potentially corresponding with post-glacial colonization patterns and latitudinal changes in floristic provinces. We find eastern massasauga populations in southern Michigan have the highly variable levels of inbreeding and genetic drift, with both the lowest and among the highest proportions of the genome found in runs of homozygosity. In general, eastern massasauga population outside of Michigan show signs of elevated inbreeding and genetic drift compared to Michigan populations. I find that the duration and extent of disturbance of natural land cover around of eastern massasauga populations is a predictor of contemporary levels of inbreeding and drift in the genome. This suggests that land use conversion has resulted in eastern massasauga populations that are smaller and more isolated compared to historical levels.

In **Chapter 4**, I explore patterns of inbreeding from over 1000 individuals from two populations of eastern massasaugas using genomic data, over a decade of capture-recapture data, and reconstructed empirical pedigrees; using these data, I test for inbreeding depression using two primary components of fitness: survival and reproduction. I find that while most individuals have low levels of inbreeding, individuals that have elevated inbreeding have lower apparent annual survival and lower probability of having pedigree offspring. This is the first documented evidence of inbreeding depression in rattlesnakes. Within each population, related individuals, like dam-offspring and sire-offspring pairs, are found closer together in geographic space than unrelated pairs of individuals, suggesting that inbreeding is generated by spatial kin clustering, which results in inbreeding depression.

Widespread anthropogenic impacts on natural populations necessitates the creative use of data to improve our understanding of the trajectories of small populations and to craft effective conservation management. Genomic data combined with life history information, such as spatial capture locations, and long-term demographic data can provide great power to understand small population problems. My dissertation highlights the potential in the incorporation of individual age (**Chapter 2**) and life history (**Chapters 3 and 4**) in genomic analyses to better understand and inform the management of small populations.

## CHAPTER TWO:

Pitfalls and windfalls of detecting demographic declines using population genetics in long-lived species

### **Abstract**

Detecting recent demographic changes is a crucial component of species conservation and management, as many natural populations face declines due to anthropogenic habitat alteration and climate change. Genetic methods allow researchers to detect changes in effective population size ( $N_e$ ) from sampling at a single timepoint. However, in species with long lifespans, there is a lag between the start of a decline in a population and the resulting decrease in genetic diversity. This lag slows the rate at which diversity is lost, and therefore makes it difficult to detect recent declines using genetic data. However, the genomes of old individuals can provide a window into the past, and can be compared to those of younger individuals, a contrast that may help reveal recent demographic declines. To test whether comparing the genomes of young and old individuals can help infer recent demographic bottlenecks, we use forward-time, individual-based simulations with varying mean individual lifespans and extents of generational overlap. We find that age information can be used to aid in the detection of demographic declines when the decline has been severe. When average lifespan is long, comparing young and old individuals from a single timepoint has greater power to detect a recent (within the last 50 years) bottleneck event than comparing individuals sampled at different points in time. Our results demonstrate how longevity and generational overlap can be both a hindrance and a boon to detecting recent demographic declines from population genomic data.

## Introduction

Modern-day wild populations are experiencing declines and bottleneck events from an unprecedented array of threats, including climate change, habitat fragmentation, and disease (Haddad et al., 2015; McCarty, 2001; Pimm et al., 2014). Population responses to these threats are often species-specific and context-dependent (Debinski & Holt, 2000; Fitzpatrick et al., 2023); for example, populations of species with specialized habitat requirements are more likely to experience sharp demographic declines when their habitat is disturbed than populations of generalist species. Populations that have undergone dramatic declines are at risk of extirpation due to environmental and demographic stochasticity and have a higher risk of experiencing inbreeding depression compared to populations that have been historically small (D. Charlesworth & Willis, 2009; Lande, 1988). The timely detection of recent population declines is essential for effective conservation intervention and ensuring the persistence of imperiled populations.

Although critical, detecting recent changes (within ~50 years) in population size is a long-standing challenge in conservation biology and evolutionary theory (Peery et al., 2012). Field-based methods to detect population size changes (Chapman, 1954; Gazey & Staley, 1986), such as mark-recapture models and long-term population monitoring, can be used to generate precise estimates of census population size ( $N_c$ ) and other demographic trends over time (Congdon et al., 1993). However, estimation of  $N_c$  from field methods is labor-intensive, difficult to conduct for cryptic species, and cannot provide estimates of historic population size.

Genetic methods to detect changes in population size are often faster and logistically easier than field-based methods, and become more affordable and informative every year (Hohenlohe et al., 2021). Rather than measure  $N_c$  directly, these methods estimate changes in

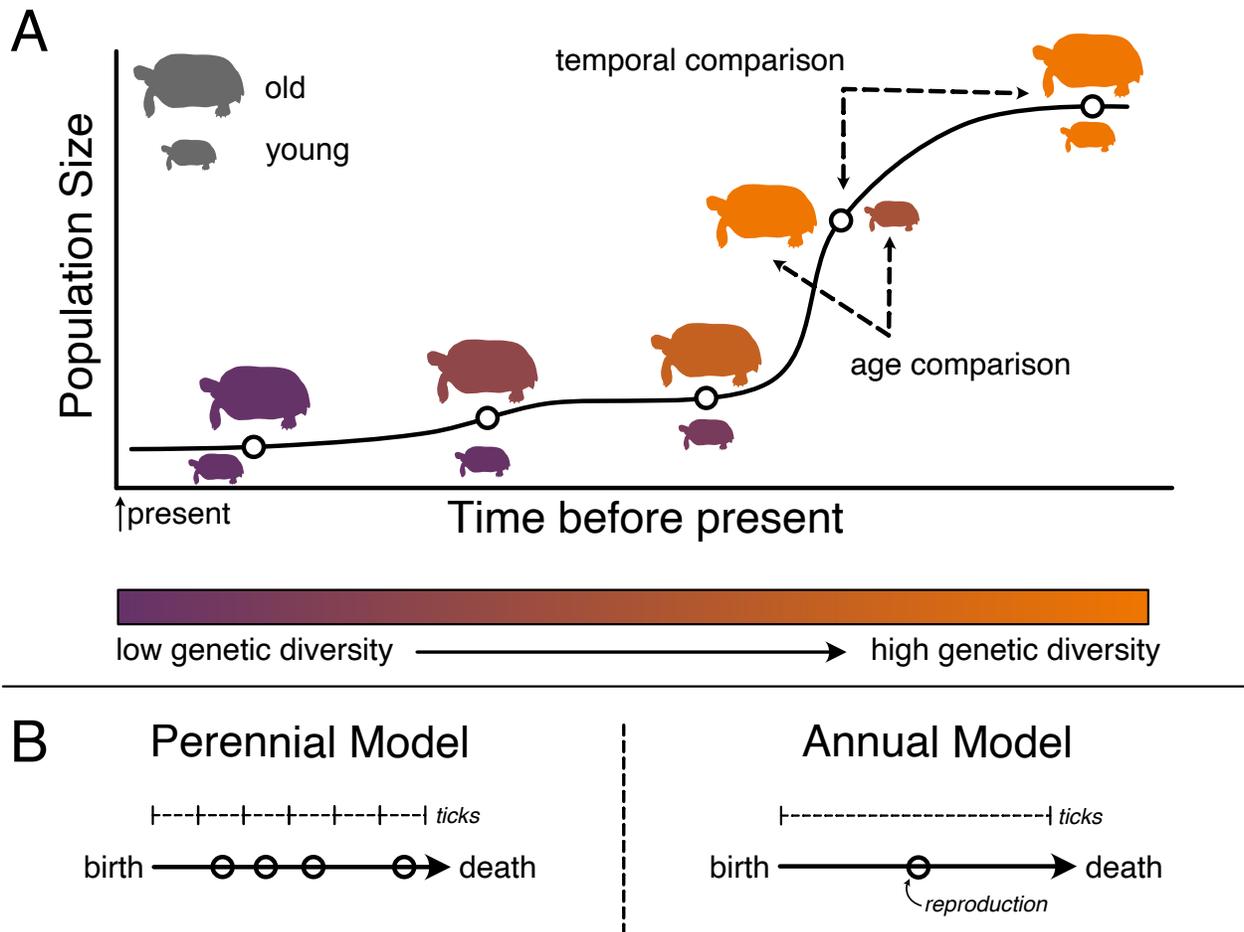
effective population size ( $N_e$ ) over time from genetic data, often from a single sampling point.  $N_e$  represents the number of individuals in an idealized Wright-Fisher population with the same genetic behavior (e.g., same rate of genetic drift, or same amount of genetic diversity) as the focal empirical population (Charlesworth, 2009; Waples, 2022). Because of many factors, including changes in population size and the magnitude of heterogeneity in reproductive success,  $N_e$  is usually smaller than  $N_c$  for a given empirical population (Fisher, 1923, p. 19; Frankham, 1995; Nunney, 1993; Turner et al., 2006; R. S. Waples, 2024, p. 202; Wright, 1931). Current methods of detecting demographic trends from genetic data are quite accurate when inferring ancient demographic trends ( $> \sim 100$  generations in the past) but have low power to capture trends associated with recent events (within the last  $\sim 100$  generations) without large amounts of sequencing data and/or individuals sampled (Antao et al., 2011; Beichman et al., 2018; R. D. Clark et al., 2023; Reid & Pinsky, 2022). Repeated temporal sampling can be a powerful tool for detecting genetic patterns associated with a decline, especially when sample size is limited (Antao et al., 2011; R. D. Clark et al., 2023), but such sampling is unlikely to exist for most populations of concern.

Using genetic methods to detect demographic declines is particularly fraught in species with long lifespans and overlapping generations because the loss of genetic diversity associated with a decline happens on a timescale determined in part by generation time (Felsenstein, 1971; Gargiulo et al., 2024; Hill, 1972; Kuo & Janzen, 2003). Generational overlap slows the rate of genetic drift; individuals contribute to multiple future cohorts, slowing the loss of genetic variation compared to organisms with less generational overlap. Long lifespan also reduces the efficacy of temporal sampling because more years will have to elapse before an appreciable change in  $N_e$ , as measured by genetic diversity, is detected. Generational overlap and long

lifespan, which are usually phylogenetically correlated (Jones et al., 2018), generate a lag between a demographic decline and the resulting decrease in  $N_e$  and genetic diversity (Nunney, 1993). This phenomenon reduces the power of demographic inference and bottleneck tests to detect recent population declines in long-lived species. In multiple empirical systems of long-lived species, including the copper redhorse (Lippé et al., 2006), ivory gull (Charbonnel et al., 2022), spotted turtles (Davy & Murphy, 2014), black abalone (Wooldridge et al., 2024), and white-tailed eagles (Hailer et al., 2006), researchers have had difficulty detecting population bottlenecks with genetic data despite strongly suspected or documented decline in  $N_e$ . Further, simulation studies have found lower accuracy in detecting demographic declines with increased generation time (Bradke et al., 2021; Reid & Pinsky, 2022).

The longevity of a species can represent a hurdle for detecting demographic declines with genetic data, but it also presents an opportunity. First, generational overlap should decrease the rate at which variants are lost due to genetic drift after a population decline, giving conservation practitioners more time to intervene before inbreeding depression and loss of adaptive potential send a population into the extinction vortex (Gilpin & Soulé, 1986). The timing of decline and extinction will be nuanced and will depend on the life history of the organism. However, the negative fitness consequences associated with inbreeding can be generated in as little one generation of non-random mating (Conner & Hartl, 2004). Classic empirical work on experimental populations of *Clarkia pulchella* found evidence of reduced fitness and increased probability of extinction in populations with small  $N_e$  after three generations (Newman & Pilson, 1997). Second, surviving individuals in the post-decline population retain genetic diversity representative of the pre-decline gene pool, whereas younger individuals born into the post-decline population come from a smaller set of possible parents and may have lower genetic

diversity (Figure 2.1A), indicating that a relationship between individual age and genetic diversity might be an early indicator of a population decline.



**Figure 2.1.** (A) In species with overlapping generations, older individuals born before the bottleneck will have genomes representative of a larger population size. Figure shows a population of tortoises that have undergone a population decline. At sampling points (open circles) before the decline, old and young turtles have similar levels of genetic diversity. During the decline, old individuals who were born before the decline have high levels of diversity compared to younger individuals, born after the decline from a limited set of possible parents. At some point after the decline, the population reaches a new equilibrium level of diversity, and old and young individuals have similar levels of genetic diversity. Dotted lines represent comparisons presented in this study: comparisons between adjacent timepoints (temporal) and comparisons between different age bins (age). (B) Schematics depicting life cycles in the annual and perennial models. In the annual model, each individual lives for one tick of the simulation and reproduction occurs once. In the perennial model, each individual lives for multiple ticks, with a probability of mortality ( $p$ ) in each tick and has opportunities to reproduce in each tick.

Several empirical studies have compared genetic diversity between older and younger individuals in long-lived species to look for evidence of population decline or the genetic consequences of known decline and fragmentation, with mixed results (Kettle et al., 2007; Labonne et al., 2016; Pereira et al., 2023; Schmidt et al., 2018; Vranckx et al., 2012; Yiniger et al., 2014). A meta-analysis by Vranckx et al. (2012) found that studies comparing genetic diversity between adult and sapling trees in fragmented habitat overall found higher allelic richness in adult trees. On the other hand, Schmidt et al. (2018) did not find a relationship between heterozygosity and age in the long-lived Australian lungfish (*Neoceratodus forsteri*), despite habitat degradation and suspected decline. The development of new methods for estimating individual age non-invasively, for example using telomere length (Hausmann & Vleck, 2002), DNA methylation levels (Nakamura et al., 2023), and radiocarbon dating (Fallon et al., 2019), expands the applicability of age comparisons in long-lived organisms of high conservation concern. However, to understand the utility of using genetic differences and age structure within contemporary populations as evidence of demographic decline, we need to investigate the relationship between age and genetic patterns in a system where the true timing and severity of the decline is known.

In this study, we use simulations to ask under what demographic and life history conditions it is useful to treat the genomes of older individuals as pseudo-temporal sampling compared with those of younger individuals to learn about recent changes in  $N_e$ . To fully understand the genetic dynamics involved in population declines in long-lived species, we track the fate of simulated individuals in populations implemented in SLiM (Haller & Messer, 2019). We simulate organisms with varying longevity, characterize the genetic patterns associated with recent demographic changes and test if sampling from different age groups can successfully

identify recent bottlenecks of varying strengths. We compare these results to temporal sampling that does not consider age information and to simulations without overlapping generations to evaluate how longevity can be both a hindrance and a boon in detecting recent demographic declines from population genomic data. Our simulations reveal that comparisons between age groups do have power to detect severe bottlenecks in species with long lifespans, but have limited power when the bottleneck is less severe, highlighting how, and under what conditions, age information can be leveraged to learn about recent changes in  $N_e$ .

## **Materials and Methods**

### ***Simulation set-up***

Demographic scenarios were modeled using simulations tracking the fate of individuals in a population in SLiM v.3.7.1 (Haller & Messer, 2019). SLiM is a forward-time genetic simulation program that allows user to define custom genomic, evolutionary, and ecological dynamics, thus allowing us to simulate populations without Wright-Fisher assumptions (Wright, 1931). Because the Wright-Fisher model assumes non-overlapping generations and no age structure, it is unsuitable to ask questions about long-lived organisms. Using SLiM, we built two neutral non-Wright-Fisher simulations, one representing a species with a long lifespan (average age in population  $\geq 2$  simulation ticks) and overlapping generations (Figure 2.1B, “perennial simulation”), and one representing a species with a short lifespan with non-overlapping generations (Figure 2.1B, “annual simulation”). In both simulations, we simulated a single chromosome  $10^8$  bp in length, with a constant recombination rate of  $1e^{-8}$  per-base position, per gamete. Simulations contained instructions for reproduction and mortality events (detailed below), which occur at each time-step, or tick, of the simulation. In our simulation, ticks can be thought of as years. All individuals in both models were simultaneous hermaphrodites and

selfing was prohibited. Simulated populations were panmictic and had no spatial component. Each simulation is detailed below.

### ***Perennial simulation***

The perennial simulation mimics a simplified life history of species such as trees and turtles—organisms that reproduce on an annual cycle and have considerable generational overlap with long lifespans. In the perennial simulation, individuals survive for multiple reproductive ticks, with  $p$  representing the probability of mortality in each tick. The average lifespan of individuals in the model is  $\frac{1-p}{p}$ , and can be manipulated in a predictable manner by changing the value of  $p$ . Individuals have a constant mortality probability across all age classes. This produces a relatively stable age structure within a simulation, with many young individuals, and relatively fewer individuals reaching older ages.

The census population size ( $N_c$ ) is an emergent property of the simulation and can be modified using the parameters  $\hat{N}$ , our desired population size, and  $K$ , the “carrying capacity” of the model. Carrying capacity  $K$  is defined as  $\hat{N} \times \frac{1-p}{1-2p}$ . A static value of  $1 - p$  determines if an individual lives or dies within a tick of the simulation. In each tick of the model, offspring are produced from individuals who are one tick old and older. The probability of a focal individual having offspring is defined as either zero or  $1 - \frac{N_c}{K}$ , whichever is larger. If  $N_c$  is less than  $K$ , the probability of reproduction will be greater than zero and if  $N_c$  is more than  $K$ , the probability of reproduction is zero. Thus, population size is maintained at approximately  $\hat{N}$  through reproduction. A second parent individual is drawn at random from the pool of individuals one tick old and older, and a single offspring is generated per pair. The probability of having offspring does not change with age. The perennial model implements an instantaneous

bottleneck event by changing  $K$  to  $B$ , where  $B = \frac{K}{R}$  and  $R$  is an integer representing bottleneck severity, when calculating the probability of offspring, slowing reproduction, and reducing  $N_c$  after the bottleneck. The probability of mortality for all individuals is temporarily changed to  $1 - \frac{K-B}{K}$  in the bottleneck tick to ensure the population did not become extinct.

### ***Annual simulation***

The annual simulation mimics the life history of organisms such as univoltine mosquitoes, butterflies, and annual plants—organisms that have a single reproductive opportunity before dying. In the annual simulation, individuals survive for a single reproductive cycle. Census population size ( $N_c$ ) is still an emergent property of the simulation, modified using  $\hat{N}$  and  $K$ . In this simulation,  $K$  is equal to  $\hat{N} + 1$  to maintain  $N_c$  at  $\hat{N}$ .

In each tick of the simulation, offspring are produced from the previous generation of individuals (all one tick of age). The number of offspring produced per focal individual is drawn from a Poisson distribution with a lambda equal to the probability of producing offspring,  $\frac{K}{N_c}$ . This probability can deviate from 1 ( $K = N_c$ ) because  $N_c$  fluctuates stochastically around  $\hat{N}$  over time. The model implements a bottleneck event by changing  $K$  to  $B$  where  $B = \frac{K}{R}$  (where  $R$  again represents bottleneck severity) when calculating the probability of offspring.

### ***Simulation implementation and analysis***

All perennial simulations ran for a burn-in period of approx.  $2 \times N_e \times \frac{p}{1-p}$  ticks (Table 2.1) to ensure equilibrium age distributions. We estimated pre-bottleneck  $N_c$  as  $N_e = \frac{\theta_W}{4\mu}$ , where  $\theta_W$  is Wu and Watterson's theta (Watterson, 1975). The annual simulations were run with a burn-in of 100,000 ticks. After the burn-in period, a bottleneck event occurred, where  $N_c$  was reduced by a factor  $R$ . All simulations were run for 500 ticks after the bottleneck event. SLiM

recorded genealogical information for 500 randomly selected individuals every 50 ticks for 200 ticks before the bottleneck, every 5 ticks for 50 ticks after the bottleneck, and every 50 ticks for an additional 450 ticks until the end of the simulation for a total of 24 sampling timepoints. We recorded age (in ticks) for every sampled individual. At the end of the simulations, we output the genealogical history of sampled individuals as tree sequences, a computationally efficient file format for storing information about genealogical networks and genetic variation within a sample (Kelleher et al., 2018; Lewanski et al., 2024).

We ran 100 replicates of the annual simulation for each of three values of  $R$  (2, 10 or 100). We ran 100 replicates of the perennial simulation with four different values of  $p$ , and three different values of  $R$  (Table 2.1, 1200 simulations total). Overlapping generations decreases  $N_e$  more than expected compared to Wright-Fisher expectations because fewer individuals are born into the population each reproductive cycle (assuming constant population size), and there is higher variance in lifetime reproductive success (Hill, 1972). To achieve similar pre-bottleneck  $N_e$  between simulations, we raised  $N_e$  in the perennial models such that  $N_e$  of each model was equivalent (Annual model  $N_e$ : 25,399, see Table 2.1 for  $N_e$  of perennial models, Hill, 1979).

### ***Tree sequence processing***

To efficiently analyze all simulations, we used tree sequences from SLiM to generate genetic data for downstream analyses. For tree sequences with multiple roots in which coalescence had not yet occurred, we simulated coalescence backwards-in-time (recapitation) using msprime v.1.0.2 (Baumdicker et al., 2022; Kelleher et al., 2016) and pyslim v.1.0.4 (Haller et al., 2019). To account for generation times greater than 1 during recapitation, we scaled  $N_e$  and recombination rate by generation time. Recapitation also ensured that simulations reached equilibrium levels of genetic diversity. We then overlaid neutral mutations on the tree sequence

using pyslim with a mutation rate of  $1e^{-8}$ . Although generation time is correlated with mutation rate in empirical systems (Bergeron et al., 2023), we chose a fixed mutation rate for all simulations to better assess the impact of lifespan alone on genetic patterns. For perennial models, the mutation rate was divided by the average generation time (as in Battey et al., 2020). To calculate generation time, we ran simplified versions of the perennial model for ten replicates per  $p$  value, with no mutations or bottleneck events, and recorded the average age of parents. Generation time was approximately  $\frac{1-p}{p} + 1$  (Table 2.1), reflecting that individuals did not reproduce until one cycle of age.

### ***Detecting declines***

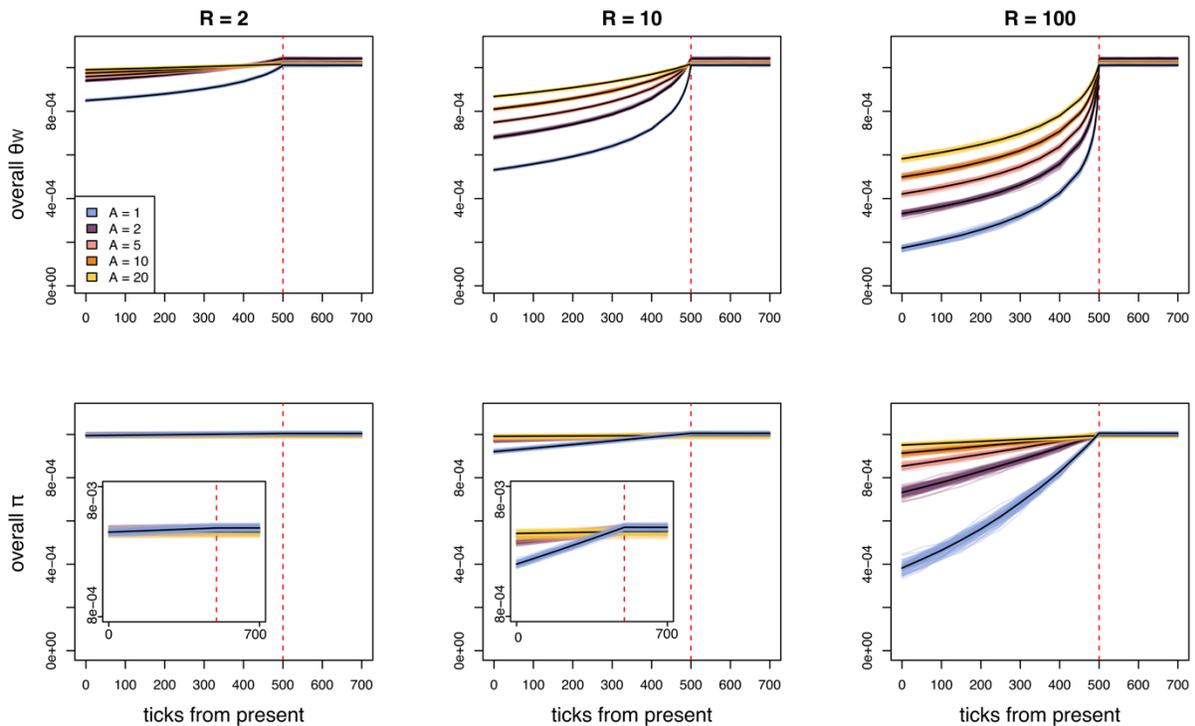
There are a variety of bottleneck detection methods available for reduced-representation and whole genome-sequencing data. In this paper, we focus on exploring the use of age information in improving detection, rather than testing commonly used modeling approaches that use genetic data to reconstruct the demographic history of populations, which enabled us to test a wider array of parameters within computational constraints. We compare age sampling from within a single timepoint to temporal sampling using two commonly used summary statistics: Wu and Watterson's theta, ( $\theta_W$ ) and pi ( $\pi$ ).  $\theta_W$  is based on the number of segregating sites in the sample, a metric that is sensitive to changes in population size and is generally comparable across groups with different sampling sizes (Peery et al., 2012; Watterson, 1975).  $\pi$  or nucleotide diversity, is based on pairwise nucleotide differences between a set of samples (Nei & Li, 1979). Because segregating sites are lost due to genetic drift during a bottleneck event, we expect  $\theta_W$  to be more reactive to simulated bottlenecks compared to  $\pi$ . Both  $\theta_W$  and  $\pi$  represent the amount of genetic diversity in a population, and directly reflect long-term  $N_e$  (R. S. Waples, 2022).

Using tree sequences for all parameter combinations and replicates, we calculated  $\theta_W$  and

$\pi$  at each sampling timepoint using four sampling schemes: (1) all individuals sampled to track trends in genetic diversity over time ( $n = 500$ ), and (2) subsamples (“bins”) containing the oldest and youngest individuals in the sample (Figure 2.1A). We defined bins based on individuals whose age at a given timepoint fell above the 90<sup>th</sup> quantile (“old bin”), and then selected the same number of the youngest individuals (“young bin”) for a given combination of simulation parameters and timepoints. Sample sizes ranged between 102 and 132 individuals (Table S2.1). (3) We repeat binned sampling such that the “old bin” contained randomly sampled individuals whose age fell above the 50<sup>th</sup> quantile, keeping sample size consistent with the previous sampling scheme. The binned sampling schemes represents a scenario where exact age information is not available, but individuals are able to be roughly categorized based on age, with varying degrees of certainty. Finally, (4) individuals selected at random from the sample, regardless of age, with the same sample size as binned sampling, to be used for temporal tests (Figure 2.1A).

With imperfect age information, we asked if the difference in a genetic summary statistic between age bins is significant. To assess significance, we used a non-parametric permutation approach. At each timepoint, sampled individual ages were randomly shuffled, and individuals were then binned according to the binned sampling schemes described above. We calculated  $\pi$  and  $\theta_W$  for each permuted bin for 100 permutations and calculated the difference between the two groups for each permutation to create a distribution of differences. We then calculated the one-tailed 95% quantile of that distribution and recorded if the observed difference between bins (sampling schemes 2 and 3) lay outside that boundary. To represent temporal sampling, we compared summary statistics between randomly sampled individuals at a timepoint to the previous timepoint (either 50 or 5 ticks in the past). We did 100 permutations where individuals

were randomly sampled from the present and past timepoints, calculated the difference in  $\theta_W$  and  $\pi$  between permutations and assessed if the observed difference (sampling scheme 4) between timepoints lay beyond the 95% quantile of the permuted distribution.



**Figure 2.2.** Plots show  $\theta_W$  and  $\pi$  from all sampled individuals across time (ticks) for the perennial (average age,  $A = 2, 5, 10, 20$ ) and annual ( $A = 1$ ) simulations with three bottleneck severities ( $R$ , bottleneck severity increases with  $R$ ). Colored lines represent different simulation replicates and black lines are the mean values across replicates. The red dashed line indicates when the bottleneck occurred.

## Results

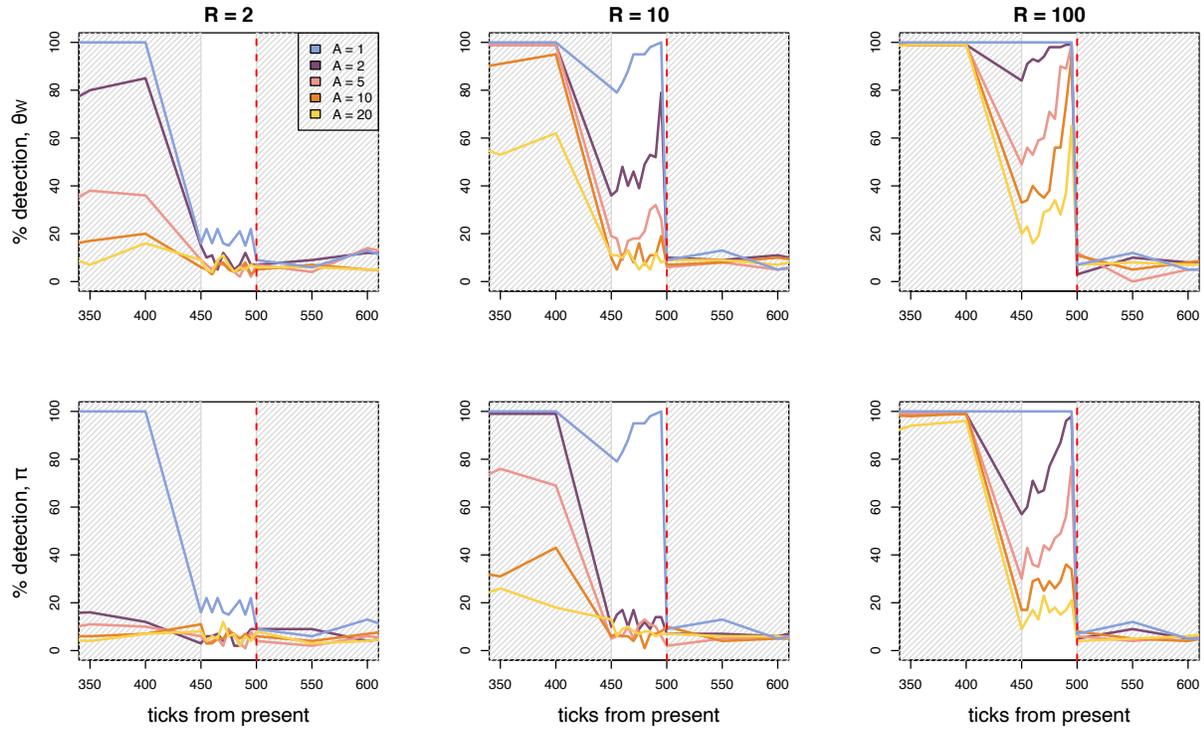
### *The impact of lifespan and overlapping generations on genetic patterns*

The rate at which genetic diversity metrics ( $\theta_W$  and  $\pi$ ) were lost after the bottleneck event decreased with increasing lifespan and corresponding generational overlap (Figure 2.2). Fifty ticks after the bottleneck decreased population size by a factor of 10,  $\theta_W$  of all sampled individuals had decreased by 21.6% when average lifespan was 1 (annual simulation), 11.7%

when average lifespan was 2, 7.2% when average lifespan was 5, 4.4% when average lifespan was 10, and 2.5% when average lifespan was 20 (Figure 2.2). As expected, overall  $\pi$  was less reactive to the bottleneck. 50 ticks after the bottleneck decreased population size by 10,  $\pi$  had decreased by 0.9% when average lifespan was 1 (annual simulation), 0.3% when average lifespan was 2, 0.1% when average lifespan was 5, 0.07% when average lifespan was 10, and 0.04% when average lifespan was 20 (Figure 2.2). As expected, we also saw that the most severe bottlenecks resulted in the fastest loss of genetic diversity. In the simulations where average lifespan was 20,  $\theta_W$  had decreased by 16.1% when the bottleneck intensity was 100, 2.5% when the bottleneck intensity was 10, and 0.3% when the bottleneck intensity was 2 by 50 ticks after the bottleneck.

### ***Temporal sampling***

When comparing diversity from before and immediately after the  $R = 10$  bottleneck event, we detected a change in  $\theta_W$  in 100% of replicates when average lifespan was 1, 79% of replicates when average lifespan was 2, 26% of replicates when average lifespan was 5, 19% of replicates when average lifespan was 10, and 8% of replicates when average lifespan was 20 (Figure 2.3). Percent detection of bottlenecks increased with increasing bottleneck intensity. As expected, directly after the bottleneck, temporal comparison separated by five ticks have more power to detect the bottleneck in the annual model compared to the perennial models (Figure 2.3). Starting 50 ticks after the bottleneck, when the timepoints being compared were 50 ticks apart, temporal comparison performed similarly in annual and perennial models under the severe bottleneck scenario ( $R = 100$ ). When the bottleneck is less severe ( $R = 2, 10$ ), temporal comparisons in the longer lifespan models ( $A = 5, 10, 20$ ) have lower detection power compared to the annual model, even when timepoints are 50 ticks apart (Figure 2.3).



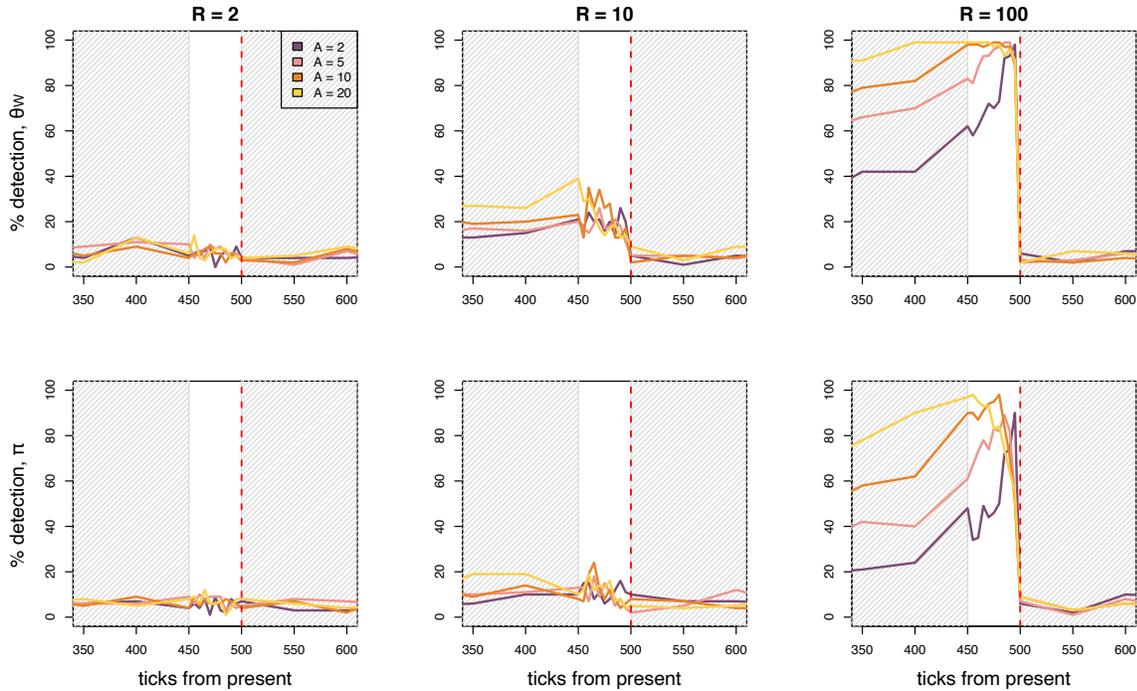
**Figure 2.3.** Plots show percent detection using temporal sampling for annual model ( $A = 1$ ) and perennial models ( $A = 2, 5, 10, 20$ ) for three bottleneck severities ( $R$ , bottleneck severity increases with  $R$ ). Percent detection is defined as the percent of replicates where we found a significant difference between individuals subsampled from the present and previous timepoints. The red dashed line indicates when the bottleneck event occurred. The gray shaded portions of the plot indicate when samples were taken every 50 ticks of the simulation, whereas the unshaded portion of the plot indicates when samples were taking every five ticks of the simulation.

### *The utility of age information*

To assess the relationship between genetic patterns and age information after a bottleneck, we evaluated the correlation between age and  $\theta_W$ , and age and  $\pi$  under a scenario where individual age can only be approximated. We compared  $\theta_W$  and  $\pi$  between the oldest and youngest individuals at each sampled timepoint and assessed significance compared to a randomly permuted distribution of differences (Figure 2.4). The strength of the bottleneck strongly influenced the differences in  $\theta_W$  and  $\pi$  between the binned age groups. When average lifespan was 20, we found a significant difference between old and young individuals in 96% ( $R$

= 100), 13% ( $R = 10$ ), and 4% ( $R = 2$ ) of replicates five ticks after the bottleneck. Fifty ticks after the bottleneck, we found significant differences in 99% ( $R = 100$ ), 26% ( $R = 10$ ), and 13% ( $R = 2$ ) of replicates for the same simulations (Figure 2.4). When we expand the older age bin to include individuals above the 50<sup>th</sup> quantile of ages, detection is similar across bottleneck severities five ticks after the bottleneck ( $R = 100$ : 95%,  $R = 10$ : 13%,  $R = 2$ : 4%), but decreases by 50 ticks after the bottleneck ( $R = 100$ : 53%,  $R = 10$ : 7%,  $R = 2$ : 9%, Figure S2.1).

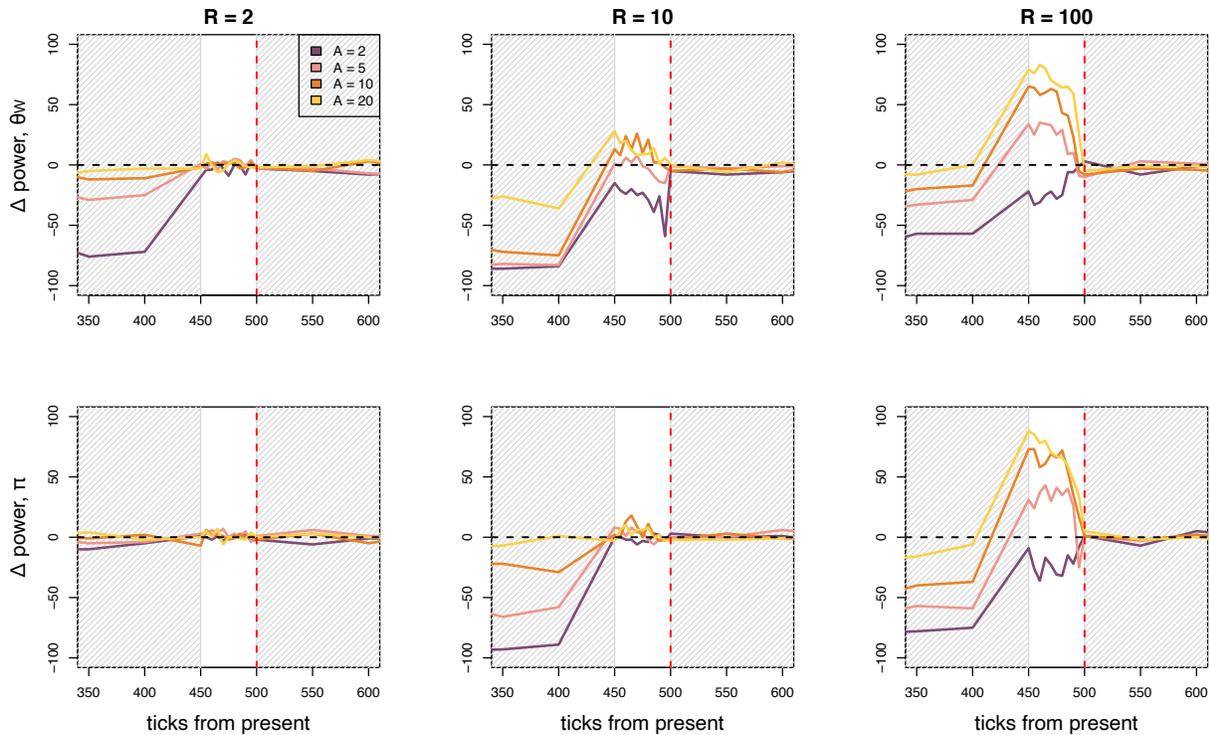
To compare the relative power of temporal sampling and age bin sampling, we calculated the difference in detection between the temporal sampling method and age bin sampling where the old bin included individuals above the 90<sup>th</sup> quantile of ages (sampling scheme 2, Figure 2.5). When the bottleneck was most severe ( $R = 100$ ), age bin sampling out-performed temporal sampling for the  $A = 5, 10$  and  $20$  simulations, but not when  $A = 2$ . Temporal sampling and age bin sampling performed similarly at lower bottleneck intensities ( $R = 2, R = 10$ ).



**Figure 2.4.** Plots show percent detection in  $\theta_W$  and  $\pi$  between old and young individuals for perennial models ( $A = 2, 5, 10, 20$ ). The old age bin is made up of individuals whose age at a

### Figure 2.4. (cont'd)

given timepoint fell above the 90<sup>th</sup> quantile. The young age bin is made up of the youngest individuals at a given timepoint. Sample size is the same between bins. Percent detection is defined as the percent of replicates where we found a significant difference between the old and young age bins. The red dashed line indicates when the bottleneck event occurred. The gray shaded portions of the plot indicate when samples were taken every 50 ticks of the simulation, where the unshaded portion of the plot indicates when samples were taking every five ticks of the simulation.



**Figure 2.5.** Plots show the difference in detection power between temporal and age bin sampling methods. Positive values indicate that age bin methods had higher percent detection and negative values indicate that temporal methods had higher percent detection. The red dashed line indicates when the bottleneck event occurred. The gray shaded portions of the plot indicate when samples were taken every 50 ticks of the simulation, where the unshaded portion of the plot indicates when samples were taking every 5 ticks of the simulation.

## Discussion

### *Use of age information in detecting declines*

Effective conservation of imperiled species necessitates the clever use and exploration of data, and our results highlight that age information could be a valuable tool when assessing the recent demography of long-lived species. In this study, we demonstrate the impact of

generational overlap and longevity on genetic diversity and test the utility of age information when detecting demographic declines in species with overlapping generations. In accordance with theory, we found that the rate at which diversity was lost after a bottleneck event decreased with increasing average lifespan in our simulations (Figure 2.2). Also as expected, we find that  $\theta_W$  is more sensitive to declines than  $\pi$  (Figure 2.2). We simulated temporal sampling by comparing diversity between adjacent timepoints and found that percent detection of the bottleneck event decreased with increasing lifespan. Temporal sampling before and after a bottleneck event is unlikely to be available for most populations. Our results suggest that age sampling—comparing younger and older individuals within a population, could support the inference of a recent bottleneck from a single sampling point, but that its use is limited when a decline has been less severe ( $R = 2, 10$ ). In general, age comparisons outperformed temporal tests when the generation time in a simulation was  $\geq$  five and when the bottleneck was severe. The change in genetic diversity after less severe declines (Figure 2.2) may not be great enough to detect with the methods used here, especially when using age sampling. We may have seen higher detection probabilities if we had used more individuals in each age bin ( $n > 130$  individuals), but these large sample sizes might be unrealistic for populations of conservation concern. Below, we discuss considerations for empirical applications of using age information when studying declines, including potential avenues to increase power.

### ***Considerations for empirical applications***

Age structure is being recognized as an important source of information about long-lived species (Dolan et al., 2023; Holmes & York, 2003). Although age information is rarely readily available for empirical populations, it is sometimes possible to approximate an individual's age from their phenotype (e.g., turtle scutes and carapace length (Jensen et al., 2018; Wilson et al.,

2003), rattlesnake rattles (Heyrend & Call, 1951), tree rings (Shroder, 1980), otoliths (Campana, 1999), tooth wear (Hinton et al., 2023), telomere length (Hausmann & Vleck, 2002), DNA methylation (Nakamura et al., 2023), coloration (Pyle, 1997) etc.) or population monitoring (Eaton & Link, 2011)). Noninvasive aging methods seem particularly promising for populations of conservation concern.

Here, we find that even having approximate age information increases power for detecting patterns of severe decline compared to temporal sampling, though increasing the uncertainty around adult age classes by expanding the older age bin decreases detection (Figure S2.1). Using individuals binned by age ranges (as opposed to exact age information) might be superior to precise age information, as small sample sizes within each age cohort ( $n \approx 1-2$  individuals) will limit inference and binning individuals based on approximate age allows for comparison between larger groups.

Based on our simulations, the ideal scenario to incorporate age information into analyses would be in populations of species with a generation time greater than 11 years, where there is reason to believe a significant demographic decline has occurred, where methods are available to approximate age, ideally including estimates of age into adulthood, and where adequate sample sizes are available. Realistically, these conditions might be met in empirical populations of trees, turtles, fish, or other species with high potential longevity and physical indicators of age. Our results suggest that a relationship between age and genetic diversity is detectable under severe bottleneck scenarios ( $R = 100$ ) even when the generation time is as short as 3 years (Figure 2.4). Given our low detection probabilities using age comparisons under less severe bottleneck intensities ( $R = 2, 10$ ), our results also highlight that a lack of a relationship between age and diversity in empirical data is not necessarily evidence that no decline has occurred, which could

explain studies that have not found a relationship between age and genetic diversity in putatively declining populations (Schmidt et al., 2018). We also observe a ~5% false positive rate in our analyses, as would be expected using a 95% quantile cut-off when performing many statistical tests.

Simulating data allowed us to test questions about age and genetic patterns in a controlled fashion. We had access to error-free whole-chromosome sequences for our samples and were only constrained by computational limitations. In wild populations, DNA quality, type of sequencing, and data quality will impact the power of inference. We expect that power to detect a relationship between age and genetic diversity will scale with the number of markers in the dataset.

To make our simulations computationally tractable, we made several simplifying assumptions in the perennial models. First, we assumed a constant probability of mortality across age classes. Many empirical populations experience higher mortality at younger ages, and lower mortality in adults (Cook, 1979; Pike et al., 2008). Increased mortality of younger individuals will increase the fraction of the population that never reproduces, limiting the pool of possible parents in future generations and increasing the rate of genetic drift. Additionally, we allowed individuals in our perennial simulations to produce a maximum of two offspring per tick. Empirical populations with increased variance in family size will have decreased  $N_e$  compared to our simulations (Hill 1979). Thus, empirical populations likely experience more genetic drift compared to our simulation populations, which could further accentuate differences between young and old individuals. We also assumed all individuals were hermaphroditic, a common simplifying assumption in evolutionary modeling.

We simulated populations isolated from the effects of gene flow. Immigrants from

populations with different demographic trajectories are likely to complicate comparisons between older and younger individuals. Additionally, we simulated a bottleneck event that occurred instantaneously. Empirical examples of such declines include natural disasters and disease outbreaks (Hsu et al., 2017; McCallum, 2012), however many actual declines are gradual and occur on the timescale of decades (although may also be of much higher magnitude). Future research should investigate the use of age sampling when declines occur over a longer period.

Selection dynamics too will alter patterns seen in empirical systems. Inbreeding depression can result in viability selection, where adults in a population have higher heterozygosity because low-heterozygosity individuals do not reach adulthood due to deleterious homozygous genotypes (Clegg & Allard, 1973; Doyle et al., 2019; Labonne et al., 2016). Here we show that adults can have higher genetic diversity compared to juveniles due to purely neutral processes if there has been a recent population decline. It is possible that viability selection and neutral processes after a bottleneck could have an additive effect on the diversity differences between age groups.

Many different calculations of  $N_e$  exist; these are reflective of different attributes of the focal population, or of the focal population at varying points in the past (Nadachowska-Brzyska et al., 2022). In this study, we use  $\theta_W$  and  $\pi$  as reflections of long-term  $N_e$ . Calculations of  $N_e$  more sensitive to contemporary changes in population size, like patterns of linkage disequilibrium and allele frequency changes over time, should hold even more power to detect recent declines (England et al., 2010; Hollenbeck et al., 2016; Nadachowska-Brzyska et al., 2022; R. S. Waples, 2022; R. S. Waples & Do, 2008). Although computationally infeasible within our experimental design, if appropriate genomic data is available for the population of concern, it may be fruitful to incorporate linkage information into comparisons between age bins,

especially for less severe declines. We also note that we may have had higher power to detect declines with more advanced demographic inference methods that can incorporate overlapping generations (Kamm et al., 2020), but to our knowledge, there are no demographic inference methods that incorporate the age of sampled individuals, though this could be an intriguing avenue of future development.

### ***Life history impacts patterns of neutral genetic diversity***

Our results highlight that the life history of an organism is intrinsically linked to  $N_e$  and to neutral genetic patterns at the individual and population level. Common model assumptions, like that of Wright-Fisher models, are violated by long-lived species in ways that impact our interpretation of data, particularly when inferring changes in  $N_e$ . Because evolution in long-lived species moves at a slower pace compared to shorter lived species (Bergeron et al., 2023), they may be more at risk of being outpaced by shifting climatic conditions and landscape changes. However, these species also maintain higher  $N_e$  compared to species with shorter generation times after declines, perhaps allowing for conservation intervention to rescue populations before they enter an extinction vortex (Gilpin & Soulé, 1986). In the absence of temporal sampling, our simulations suggest that age information is an additional tool researchers can use in detecting recent demographic bottlenecks in long-lived species.

### **Acknowledgements**

All simulations and analyses were run on the High Performance Computing Center at Michigan State University's Institute for Cyber-enabled Research. I would like to thank Bob Week, Matteo Tomasini, and Zach Hancock for help and feedback on the development of the simulations, Peter Ralph for guidance using `pyslim`, Tyler Linderoth for fruitful discussion on the true nature of  $\theta$ , and members of the Bradburd, Fitzpatrick, and Winger Labs, and three

anonymous reviewers for valuable feedback on this paper. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM137919 (awarded to GB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. This is W. K. Kellogg Biological Station Contribution No. 2374. This work was published in the journal *Evolutionary Applications* (John Wiley & Sons Ltd.) and is reproduced here under a CC BY Creative Commons License.

## Tables

**Table 2.1.** Average age, probability of mortality, maximum age in ticks, generation time, census population size ( $N_c$ ), and number of burn-in ticks of perennial simulations.

Average age ( $A$ )	Probability of mortality ( $p$ )	Approximate max age (ticks)	Generation time	$N_c$	Burn-in period (ticks)
2	1/3	30	2.999165	42,397	94,700
5	1/6	50	6.006728	46,870	236,800
10	1/11	90	11.01567	48,939	473,500
20	1/21	250	20.97633	50,000	947,100

## CHAPTER THREE:

Anthropogenic land use changes associated with increased inbreeding in a threatened rattlesnake

### **Abstract**

Anthropogenic land use changes have a plethora of consequences, including the loss and fragmentation of natural habitat. Many species now exist in small populations and are at high risk of inbreeding, inbreeding depression, and extirpation. Here, we use whole genome sequencing to evaluate the genetic status of populations of threatened eastern massasauga rattlesnakes (*Sistrurus catenatus*) from throughout their range, with a focus on their range center. We find strong genetic differentiation and a pattern of isolation by distance between populations. Populations at the range center harbor more genetic diversity than range edges, though all populations have runs of homozygosity indicative of inbreeding. The range edges have more long runs of homozygosity, indicative of recent inbreeding, than most populations in the range core. We use back-cast land cover data from the past four centuries to document anthropogenic land cover changes in the area surrounding each population. We find a significant relationship between inbreeding and the amount of habitat around a population that has been converted to an anthropogenic land use type, suggesting that land use changes have had a direct contribution to inbreeding. This information gives important context to conservation managers about the status of eastern massasauga populations and how much genetic diversity remains in different part of their range.

### **Introduction**

Anthropogenic land cover changes have dramatically reshaped how organisms experience their environment. Conversion of natural habitat to anthropogenic land cover is one of the most dramatic and pervasive land use changes in the United States over the past 200 years

(X. Li et al., 2023; Sohl et al., 2016). From 1850 to 1920, cropland in the conterminous United States expanded by 114.96 million hectares, whereas forest area decreased by 40.83 million hectares during the same period (X. Li et al., 2023). This large-scale conversion has had dramatic consequences for wildlife. For many species, suitable habitat has shrunk, and existing habitat is increasingly fragmented, leading to population declines and biodiversity loss (Exposito-Alonso et al., 2022; Haddad et al., 2015).

The fate of fragmented populations is nuanced. In addition to direct causes of mortality, extinction risk in fragmented landscapes is increased by inbreeding, loss of genetic diversity, and Allee effects (Gilpin & Soulé, 1986; Keller & Waller, 2002). Some species can maintain metapopulation dynamics between habitat patches through long distance dispersal (Trakhtenbrot et al., 2005), helping to dampen loss of genetic diversity and genetic drift. However, in many cases, habitat loss and fragmentation lead to decreased gene flow between populations, increased inbreeding, and population decline. For example, mountain lions (*Puma concolor*) in the Santa Monica Mountains, California, USA have become increasingly isolated because of roads and urbanization, leading to an increase in inbreeding and a decrease in fitness (Huffmeyer et al., 2022; Vickers et al., 2015).

The rate of population decline is a key factor in assessing extirpation risk (Charlesworth & Willis, 2009; Lande, 1988). Populations that have been historically small and isolated from gene flow will have accumulated fixed and segregating deleterious alleles due to genetic drift (high drift load) but are expected to purge large-effect deleterious recessive alleles that are more likely to be exposed to selection in small populations (low inbreeding load). Large populations that have undergone a recent decline will not have as many fixed deleterious alleles (low drift load) but will have more segregating high-effect deleterious recessive alleles (high inbreeding

load) (Keller & Waller, 2002). Therefore, more conservation concern with respect to inbreeding depression is placed on populations that have experienced recent declines. Testing if habitat fragmentation has resulted in recent population declines and increased inbreeding is therefore crucial to understanding the status of imperiled populations and to inform their management. We can untangle the impact of habitat fragmentation on natural populations by leveraging whole genome sequences to inform the timing of declines, levels of inbreeding and load, and how land use affects patterns of genetic diversity and relatedness.

The eastern massasauga (*Sistrurus catenatus*) is an elusive and threatened rattlesnake native to wetlands and prairie fens of the Great Lakes region of the United States and Canada. They are listed as threatened or endangered everywhere they occur. Eastern massasaugas have specific habitat requirements that include wetlands they occupy in the summer, and upland forests where they overwinter. Many anthropogenic land use types, such as agriculture and roads, are known barriers to dispersal (Martin et al., 2023), and are also associated with population extirpation (Pomara et al., 2014). Eastern massasaugas have small home ranges, and dispersal is limited (Harvey & Weatherhead, 2006; Martin et al., 2021; Moore & Gillingham, 2006). However, male long-distance dispersal is thought to be an avenue for gene flow between populations, although how habitat loss has impacted gene flow between populations compared to historical levels is unknown.

Assessing the genetic diversity, prevalence of inbreeding, and demographic histories of extant eastern massasauga populations is essential for their management, and for understanding broader trends in how habitat loss impacts wild populations. Contemporary eastern massasauga populations are small, and while there is evidence that they have been small for a long time (Chiucchi & Gibbs, 2010), there is also evidence for recent declines along their range edge

(Sovic et al., 2019). The demographic history of populations in Michigan, the range center and where eastern massasaugas populations are thought to be larger than elsewhere, is unknown. Recent work has revealed inbreeding depression, driven by dispersal limitation, in two populations in Michigan, leading to questions about the distribution of genetic diversity and inbreeding across the range of eastern massasaugas (Clark et al. in prep). Here, we leverage range-wide whole genome sequencing to (1) document population structure and the distribution of genetic diversity across the entire range of eastern massasaugas (2) infer demographic history of populations, and (3) ask if land use changes are correlated with changes in population size. We hypothesize that the lowest inbreeding and highest genetic diversity will be found in populations at the range center, and that trends in inbreeding will be at least partially explained by variation in the amount of anthropogenic land-use surrounding a population. We also expect that populations known to be small and potentially extirpated will have high inbreeding, regardless of geography.

## **Methods**

### *Whole genome sequencing and bioinformatic pipeline*

We extracted high molecular weight DNA from 202 eastern massasauga blood samples from 16 sampling sites collected from 2015 to 2022 using Qiagen extraction kits (Redwood City, CA, USA) (Table 3.1). For ten preliminary samples, libraries were prepared and sequenced by Novogene Co. (Sacramento, CA, USA) using Illumina PE 150 sequencing. The remaining 192 samples were prepared into sequencing libraries and sequenced on a NovaSeq X using 150 bp read pairs by the University of Michigan Advanced Genomics Core. An additional 117 sequences from sites outside of Michigan were downloaded from the NCBI SRA database (Mathur et al., 2023).

We removed adapter sequences from forward and reverse reads and polyG tails with a minimum length of 10 with fastp version 0.23.4 (Chen, 2023). We dropped individuals with unusual GC content distributions indicative of sample contamination or high bacterial DNA content. Trimmed reads were aligned to an eastern massasauga reference genome (GenBank Accession: GCA\_039880765.1) using bwa mem version 0.7.17 with default parameters (H. Li, 2013). We filled in mate coordinate information between read pairs using SAMtools fixmate version 1.18. Alignments were sorted and duplicate reads removed using SAMtools sort and rmdup. We removed overlapping base pairs from read pairs using bamUtil clipOverlap version 1.0.15 using default parameters. Filtered alignments for the same individual sequenced on multiple lanes were merged using SAMtools merge.

We calculated genotype likelihoods from properly paired aligned reads using BCFtools mpileup version 1.19 (H. Li, 2011). We required a minimum mapping quality (-q) of 20 and a minimum base quality score (-Q) of 13. Genotype calls were made using BCFtools call -m with a prior (-P) of 0.002 (*pi\_geneDes*, Mathur et al. 2023). Variants were normalized using BCFtools norm. We used summaries of depth, mapping quality, and base score quality to assess the quality of our alignments and inform filtering cut-offs. We extracted SNPs that passed quality filter, excluding SNPs within indels. We dropped individuals with average depth less than 12X. We retained a site if at least 90% of individuals in a population had a minimum of 10X depth and a genotype quality score of 15. We also dropped six individuals that had high heterozygosity, large numbers of unique variant sites (singletons), many alternate reads in the mitochondrial genome, and patterns of reference/alternate read counts indicative of sample contamination.

#### *Genetic variation and population structure*

Using called genotypes, we calculated heterozygosity at variant sites using VCFtools v.

0.1.17 (Danecek et al., 2011) and genome-wide  $\pi$  using pixy v.1.2.10.beta2 (Korunes & Samuk, 2021) from all sites that passed quality filtering. We calculated eigenvectors and values for our dataset using plink v.1.9 for principal components analysis (Chang et al., 2015) after implementing a minor allele frequency cut-off of 0.05, and removing linked SNPs with a pairwise  $r^2$  threshold of 0.5, a window size of 50 variants, and a step size of 5 variants. Using the same set of pruned and maf-filtered SNPs, we used ADMIXTURE (Alexander et al., 2009) to visualize population clustering and used cross validation to inform our interpretation of the most biologically relevant  $K$ -value. To further explore population clustering, we calculated pairwise genetic distance between individuals using ngsDist (Vieira et al., 2016) and built a neighbor joining tree using fastME (Lefort et al., 2015), placing 100 bootstraps on the tree using raxml-ng (Kozlov et al., 2019). We calculated pairwise  $F_{ST}$  between populations using VCFtools. We excluded sites where we had less than 3 samples, as sample allele frequencies were likely not representative of the allele frequencies in the entire population.

#### *ROH inference*

We measured inbreeding within each population using runs of homozygosity (ROH). We called runs of homozygosity using BCFtools roh (Narasimhan et al., 2016), using a constant recombination rate of  $3.935e^{-8}$  (Schield et al., 2020) and estimating hidden Markov model parameters using the Baum-Welch algorithm with a convergence threshold of  $1e^{-10}$ . We ran BCFtools roh separately for four regions correlated with genetic demes (1: Bois Blanc Island, 2: Northern Michigan, 3: Southern Michigan, and 4: snakes from Ohio, Illinois, Pennsylvania, New York, and Ontario) using allele frequency estimates for each group. We were confident in calling a ROH if it received an average quality score of 20 or higher. We calculated  $F_{ROH}$ , the proportion of an individual's genome found in a ROH tract of any size class using custom code. For each

individual, we calculated the total number of ROH (NROH). We split runs into two length categories based on when land use conversion began across the study area (before 200 years/67 generations ago, 0.05–0.19 Mb, and after 0.19–21.6 Mb). We dated ROH of varying lengths using a recombination rate of 3.935 cM per Mb, a generation time of three, (Faust et al., 2011) and the following equation:  $g = \frac{100}{2rL}$ , where  $r$  is the recombination rate,  $L$  is the ROH tract length, and  $g$  is generations in the past (Saremi et al., 2019).

### *Demographic inference*

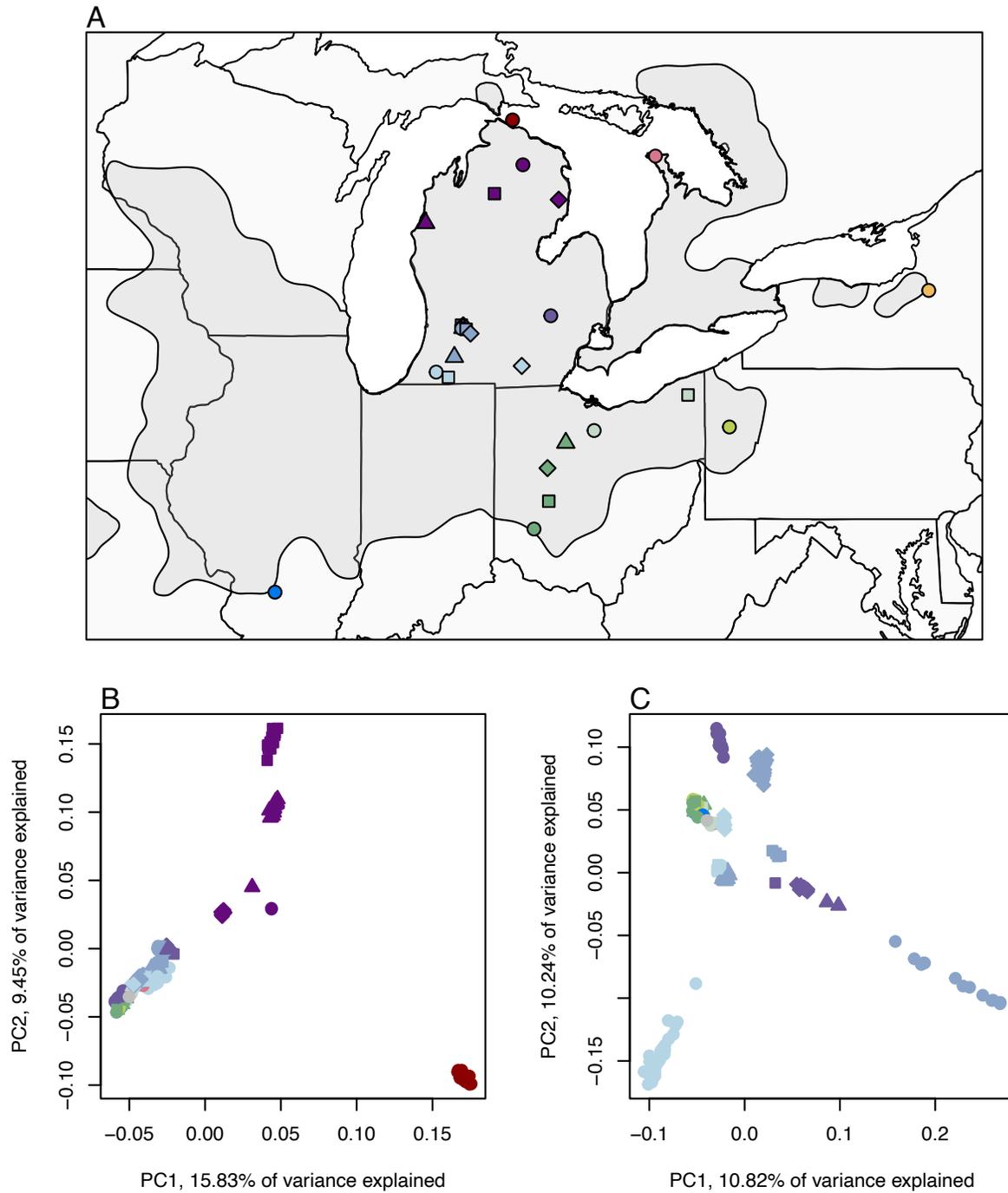
We used the program GONE (Santiago et al., 2020) to estimate changes in  $N_e$  over time. GONE uses linkage disequilibrium to infer past  $N_e$ . Because GONE performs best with larger sample sizes, we only ran this analysis on sampling sites that had 10 or more individuals. We used a constant recombination rate of 3.935 cM per Mb and set the maximum value of recombination rate to analyze ( $hc$ ) to 0.01, as recommended when focal populations have likely experienced recent migrants. Because rattlesnakes are known to have high recombination rates on their micro-chromosomes (Schield et al., 2020), we only used macrochromosomes in this analysis (as in Ochoa and Gibbs 2021).

### *Land use change*

We downloaded land use/land cover data for each year from 1630 to 2020 for the conterminous United States in 1x1 km resolution from Li et al. (2023). Each spatial cell is categorized as one of eight land cover types: “barren”, “water”, “wetland”, “grassland”, “shrub”, “forest”, “pasture”, “crop”, “urban”, or “nodata”. We used the R package “stars” in R version 4.3.2 to extract land use/land cover data from a 2-by-2 km grids centered on each sampling site. We selected a 2x2 km grid to focus on the habitat experienced by eastern massasaugas. This size is likely to encompass the home ranges ( $\sim 0.013$  km<sup>2</sup>, Moore & Gillingham, 2006) of eastern

massasaugas within each population. For each grid, we calculated the average proportion of time a cell in the grid was categorized as an anthropogenic land use type for eastern massasaugas in R. We summed values across each grid and divided by the total number of cells in the grid where land cover data was available. A value close to zero indicates that a site experienced little disturbance from 1630 to 2020. A value above zero indicates an increasing time and area within the grid occupied by anthropogenic land use types. We define potentially suitable habitat as any non-anthropogenic land cover type, and “unsuitable habitat” as pasture, crop, or urban land cover.

To evaluate the effect of disturbance on  $F_{ROH}$ , we built multiple linear regression models using the `glmmTMB` package in R v. 4.4.1. We constructed four models using two different subsets of the data. First, we modeled  $F_{ROH}$  as a function of the average proportion of anthropogenic land use at a site and genetic relatedness between individuals and populations, as approximated by PCs one through six. Second, because we expect that there are many reasons why a population in undisturbed habitat may experience inbreeding (habitat quality, long-term demographic history), we repeated our model with a subset of populations that had an average proportion of anthropogenic land use above zero. For each of these models, we also built a corresponding pared-down model that just included the PCs and no disturbance metrics. We used  $AIC_c$  (Akaike, 1973) adjusted for small sample size via the R package `AICcmodavg` (Mazerolle 2023) to determine which model was the best fit to the data.



**Figure 3.1.** (A) Map of sampling locations of eastern massasauga rattlesnakes. Shaded gray region represents the historical range for eastern (Nature Serve & IUCN (International Union for Conservation of Nature), 2007). (B) Principal components analysis of 26 sampling sites. (C) Principal components analysis of 21 sampling sites outside of Northern Michigan.

## Results

### *Whole genome sequencing*

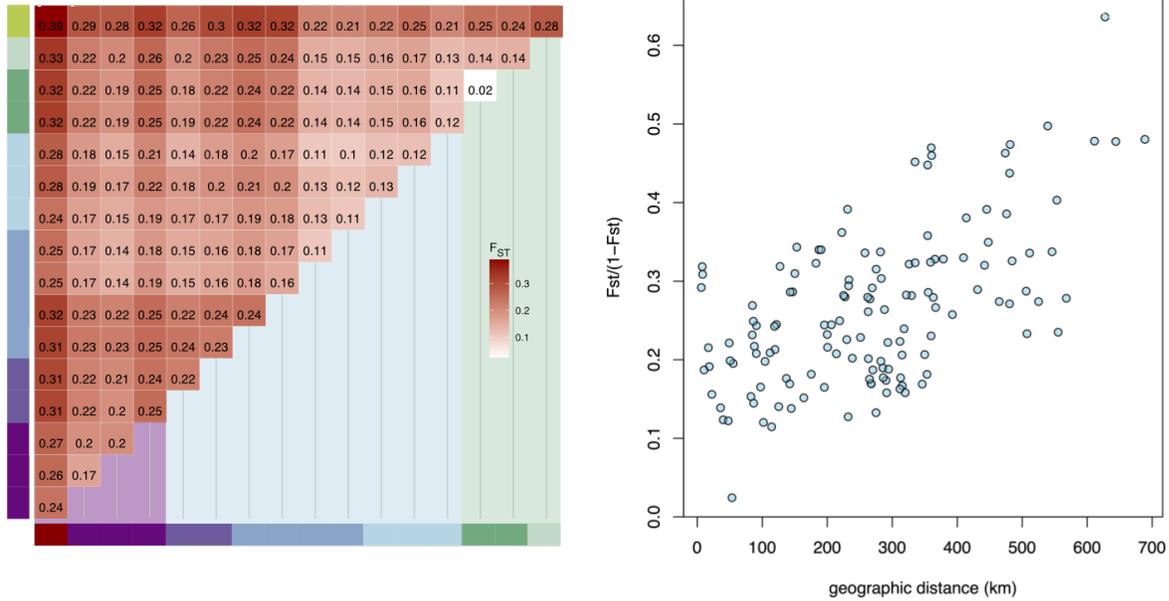
Our final data set included 2,352,466 high quality SNPs from 221 individuals from 26 sites (Figure 3.1, Table 3.1, 199 from Michigan and 22 from outside of Michigan). Our LD pruned and maf-filtered dataset included a subset of 226,628 SNPs. Average individual depth was 19.4 reads (range: 15.4-27.0). Because we combined datasets with different average number of sequencing reads, we were particularly stringent when filtering to ensure batch effect did not skew downstream analyses. The average individual depth in our final dataset for the individuals from Mathur et al. 2023 (mean: 19.4, range: 15.4-27.0) was not significantly different from new individuals (mean: 19.4, range: 15.5-26.4) ( $t(418.8) = 0.34$ ,  $p = 0.74$ ), nor was there a significant relationship between individual depth and heterozygosity at variant sites (Figure S3.1).

### *Population structure and genetic diversity*

We found high population structure between eastern massasauga populations. Principal component analysis showed that eastern massasaugas on Bois Blanc Island and in Northern Michigan are highly differentiated from other populations. Snakes in southern Michigan clustered with snakes from Ohio, Illinois, New York, Pennsylvania, and Ontario (Figure 3.1B). Our admixture results showed low amounts of admixture between populations, as each additional  $K$ -value generally separated out another sampling site as a distinct cluster (Figure S3.2). Cross-validation indicated that  $K=9$  was the most relevant  $K$ -value (Figure S3.3). All populations in Michigan and Ontario formed monophyletic groups in our neighbor-joining tree, but sites in Ohio, New York, Pennsylvania, and Illinois were paraphyletic (Figure 3.4). However, support from 100 bootstrap replicates was low for all clades.

Pairwise  $F_{ST}$  ranged from 0.024 to 0.39, with an average of 0.2 (Figure 3.2). We observed

a strong pattern of isolation by distance; genetic distance (measured as linearized  $F_{ST}$ ) is correlated with geographic distance (Figure 3.1C).



**Figure 3.2.** Heatmap shows pairwise  $F_{ST}$  values between eastern massasauga rattlesnakes from 17 populations with three or more individuals. Colors along the x and y-axes correspond to colors used in Figure 3.1 and the shaded background indicate regions (from left to right: Northern Michigan, Southern Michigan, and Ohio). Points on the right-hand plot represent a comparison between two populations ( $N \geq 3$ ) and shows a pattern of isolation by distance: a relationship between linearized  $F_{ST}$  and geographic distance (km).

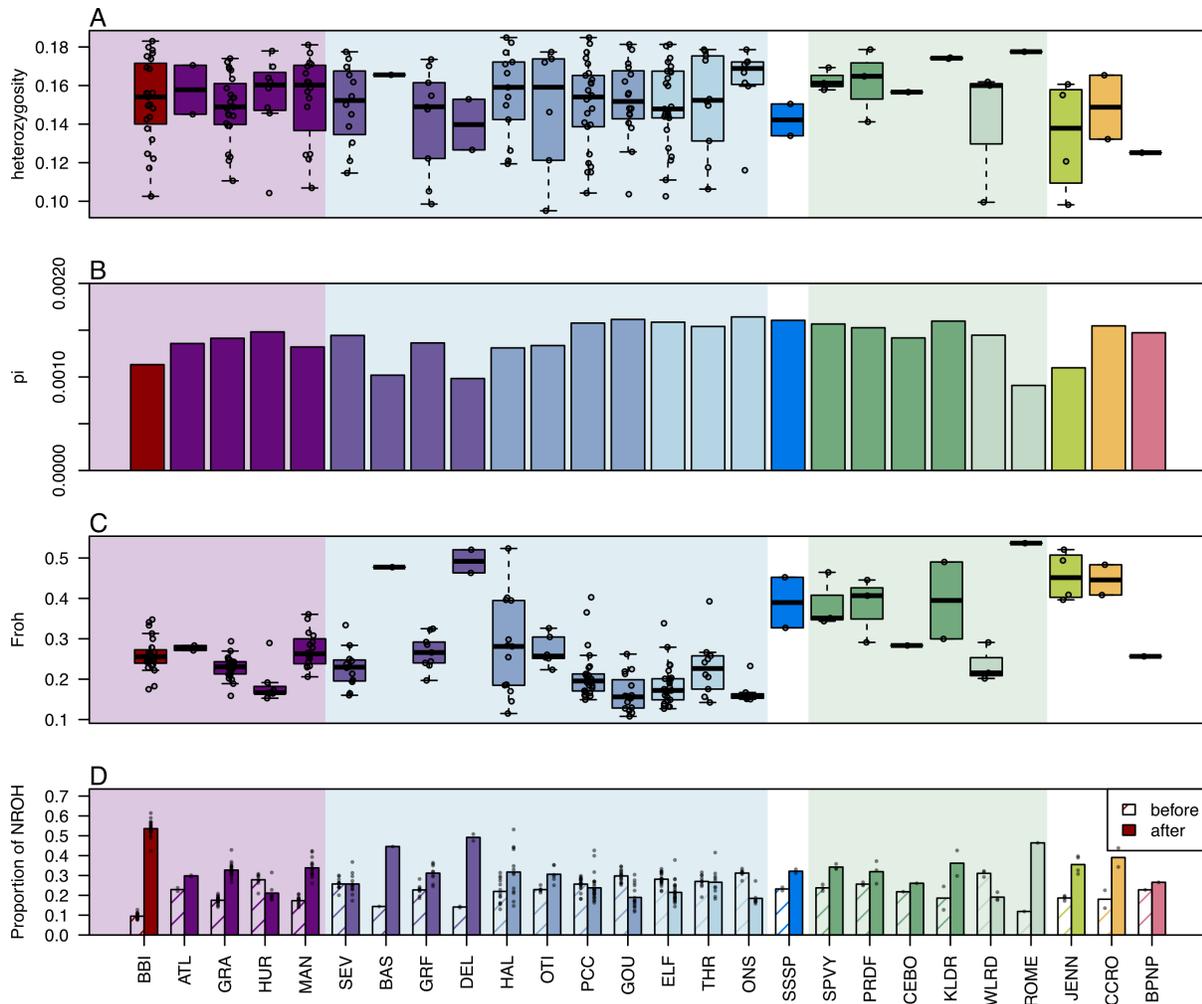
As expected, we found the lowest heterozygosity, lowest genome-wide  $\pi$ , and highest inbreeding in eastern massasauga populations at the range edges, and the highest heterozygosity and lowest inbreeding in Michigan (Figure 3.3A-C).  $F_{ROH}$ , the genomic inbreeding coefficient ranged from 0.1 to 0.5. Generally, the highest inbreeding was found in sites known to be small, or likely extirpated.

### *Runs of homozygosity*

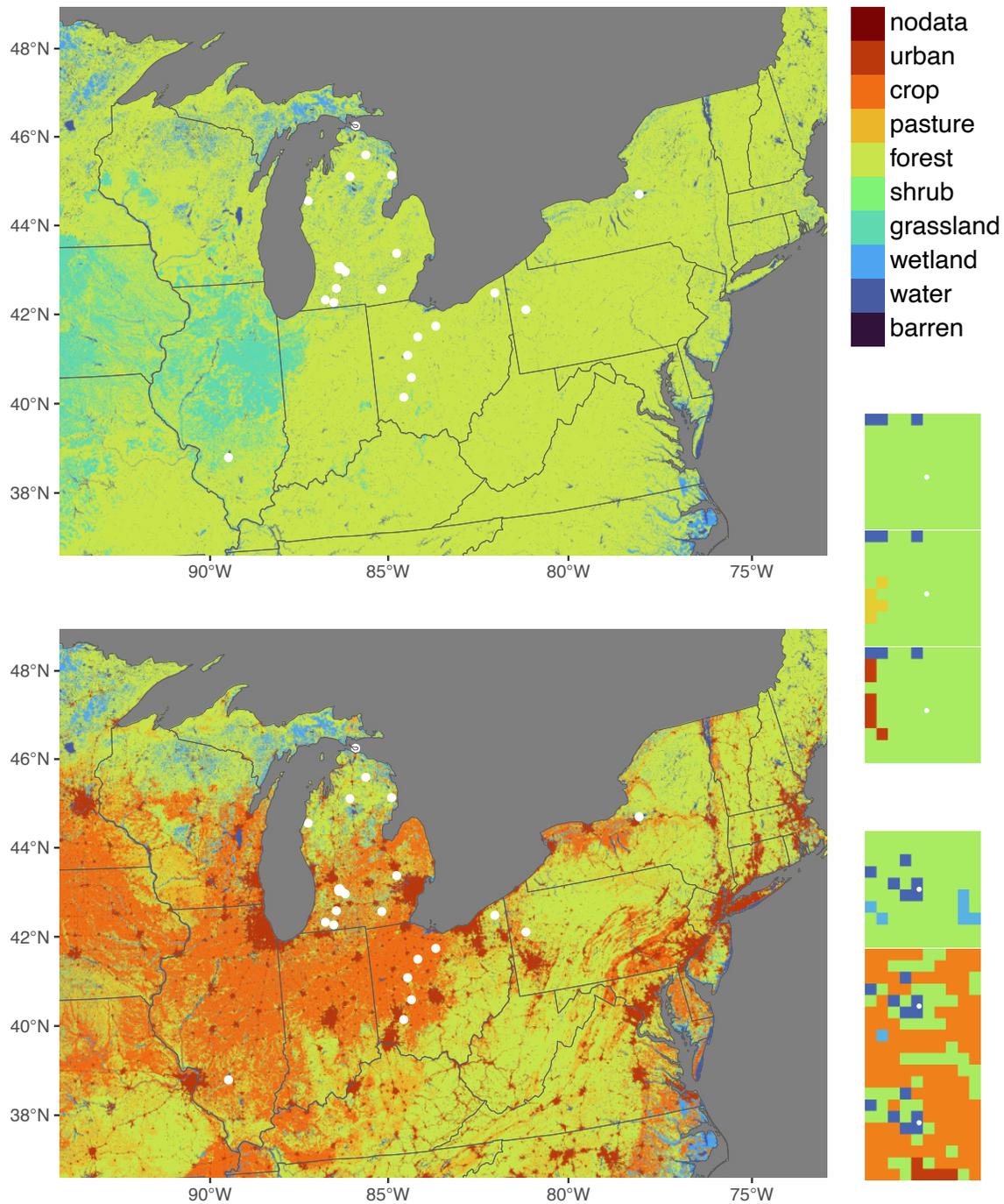
The distributions of  $F_{ROH}$  originating from before and after the beginning of widespread anthropogenic land use changes indicate differing demographic trajectories among populations. For example, sites in northern Michigan, Ohio, New York, Illinois, and Pennsylvania have

higher proportions of recent ROH compared to most sites in southern Michigan (Figure 3.3D).

Two sites in southern Michigan (“BAS”, and “DEL”) known to be small and potentially extirpated (J. Moore, *personal communication*) have elevated long ROH tracts, indicating recent inbreeding occurring at these sites.



**Figure 3.3.** (A) Heterozygosity per population of eastern massasauga rattlesnakes. Each dot represents an individual. (B) Genome-wide  $\pi$  per population (C)  $F_{ROH}$  of eastern massasauga rattlesnakes. For all boxplots, black lines represent the median at each population, and each box represents the 1<sup>st</sup> and 3<sup>rd</sup> quantiles of the data. (D) Bar plot shows the proportion of runs of homozygosity that fall into two different categories dated by length: before widespread anthropogenic land use change (0.05–0.19 Mb, 67–254 generations ago) and after (0.19–21.6 Mbp, 0.6–64 generations ago). Shaded background indicate regions (from left to right: Northern Michigan, Southern Michigan, and Ohio).

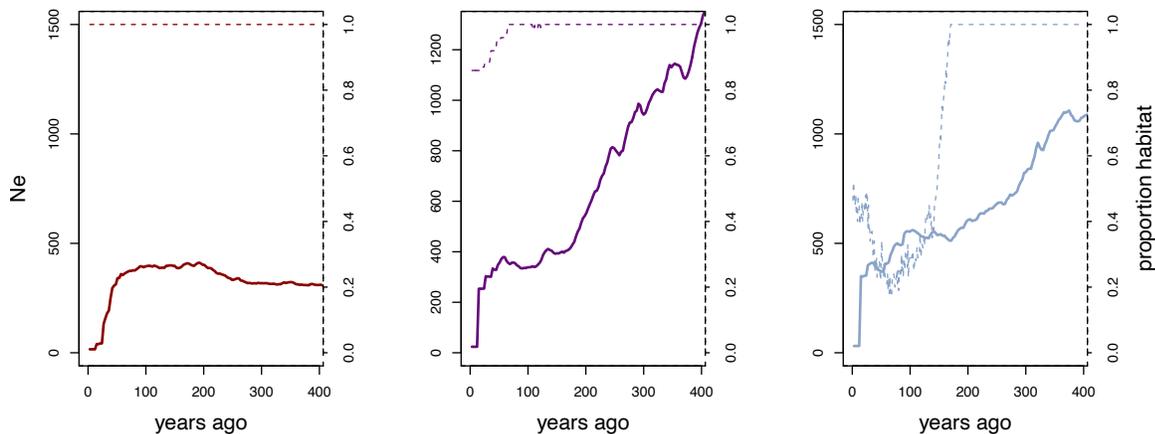


**Figure 3.4.** Figure shows maps of sampling sites overlaid on land use/land cover data from 1630 (top) and 2020 (bottom). Land use/land cover data is from Li et al. 2020. Inset shows change in land cover over time in a 10x10 grid around two focal sites (white dots). From top to bottom, squares represent land use in 1820, 1920, and 2020.

### *Demographic inference*

We used GONE to estimate  $N_e$  over the recent past. We ran GONE on eight sites that had

more than 10 individuals. The results of seven of those sites showed extreme recent population decline, preceded by a dramatic rise in  $N_e$ : patterns characteristic of model inaccuracy due to recent admixture (Novo et al. 2023). Therefore, we only present results from three populations: “BBI”, “GRA”, and “PCC” (Figure 3.5, solid lines). BBI shows a recent decline in  $N_e$  starting around 50 years before present. Population size at BBI appears stable before then. Our results suggest that  $N_e$  of GRA and PCC had been decreasing steadily until approx. 200 years ago, when  $N_e$  levels out, followed by a recent decrease starting approx. 50 years ago at GRA and 100 years ago at PCC. The declines at GRA and PCC maybe be related to decreases in the proportion of potentially suitable habitat available at each site (Figure 3.5, dashed lines).

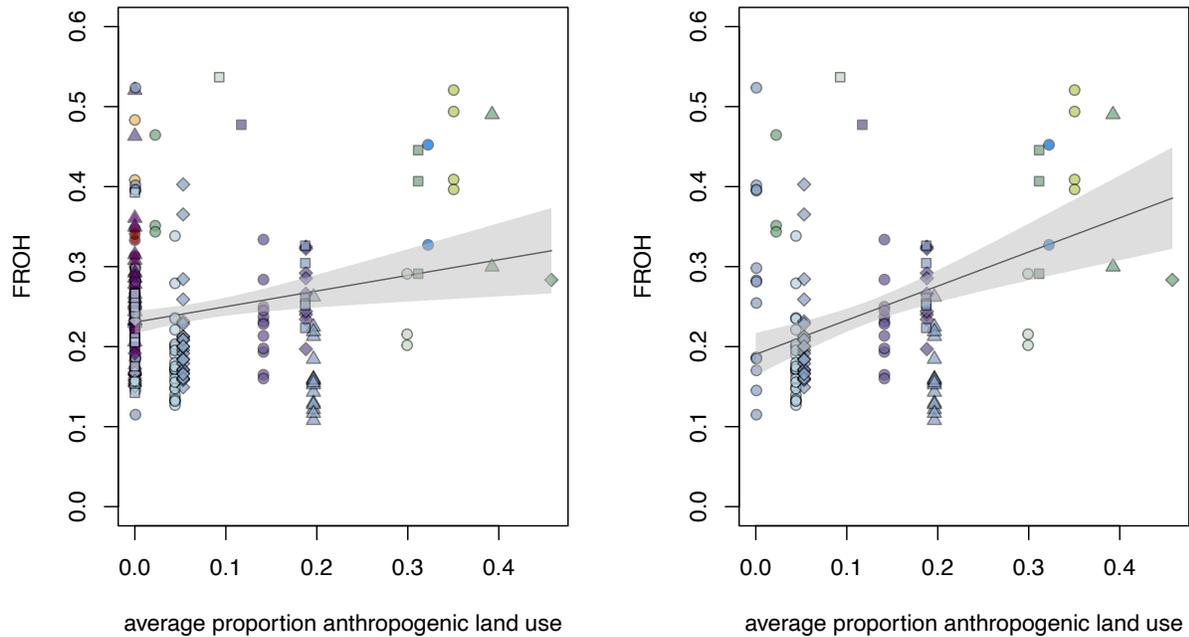


**Figure 3.5.** Effective population size ( $N_e$ ) from GONE analyses going back in time for eastern massasauga at BBI (red lines), GRA (purple lines), and PCC (blue lines). The right y-axis shows the proportion of a 10 by 10 km grid around the focal site that contained suitable habitat over time.

*Increased anthropogenic land use associated with increased inbreeding*

The amount of potentially suitable habitat lost varied by site (Figure 3.4, Figure S3.5). The most common land cover type lost within 10km of each site was forest, which was most often converted to agriculture. In 2020, the proportion of natural habitat within 10 km of each population ranged from 0.02 to 1.0. The majority of natural land cover loss occurred after 1880

(Figure S3.5).



**Figure 3.6.** Plots show the relationship between the average proportion of time a 2x2 km grid is an anthropogenic, rather than natural, land use type and  $F_{ROH}$  in eastern massasaugas from 25 populations. Values on the x-axis near 0 indicate little disturbance, while higher values indicate more long-term and/or widespread disturbance. The left plot shows the effect in the full dataset, whereas the right plot shows the effect on just sites where there was anthropogenic disturbance of habitat. Solid lines represent the predicted  $F_{ROH}$  value. The shaded band represents 95% confidence intervals.

We report a positive effect of disturbance on inbreeding in eastern massasaugas (Figure 3.6, Figure S3.6). The top-ranked models from  $AIC_c$  comparison using the full dataset and the subset of populations that experienced disturbance both included the disturbance metric (Table 3.2). Sites that have experienced more anthropogenic disturbance have elevated inbreeding compared to less disturbed sites. When considering all sites, we also found positive effects of PC1 and PC6 on inbreeding and a negative effect of PC3 and PC4 on inbreeding. When considering just sites that experienced disturbance, we found a positive effect of PC5 on inbreeding, and a negative effect of PC2 on inbreeding. The significant effects of principal components indicate varied impacts of population structure, demographic history, and

relatedness on inbreeding. The 85% confidence intervals of all other effect sizes overlapped with zero, and we interpret them as not biologically significant to our top models (Arnold, 2010).

## **Discussion**

The impact of habitat loss on wild populations is a critical concern in conservation biology. Here, we use range-wide sampling of eastern massasaugas to document population differentiation and genetic diversity and show the effect of land cover conversion on inbreeding within populations. We find that inbreeding is associated with increased anthropogenic disturbance in the landscape surrounding populations.

### *Population structure*

This is the first study which has used genomic sequencing to evaluate the relationship between eastern massasaugas in Michigan and the rest of their range. We find that eastern massasaugas in northern Michigan are highly differentiated from the rest of the range (Figure 3.1B), whereas snakes in southern Michigan are more closely related to snakes in neighboring states (Figure 3.1B). Eastern massasaugas on Bois Blanc Island are the most differentiated, likely because of their isolation from mainland populations. This pattern of differentiation could represent post-glacial colonization of Michigan from southern populations, where northern Michigan sites are genetically distinct because of successive founder events.

The region between the northern and southern Michigan sites is known as the floristic tension zone and corresponds to changes in soil composition and forest types (McCann, 1979). Many important life-history traits, like body size, litter, and offspring size (Hileman et al., 2017) and perhaps generation time (Faust et al., 2011; Miller, 2005), vary across the range of eastern massasaugas, indicating potential adaptation to different environmental conditions. The possibility of local adaptation in eastern massasaugas north and south of the floristic tension zone

should be the subject of future study.

### *Diversity and inbreeding*

In accordance with previous understanding, Michigan eastern massasauga populations, in general, have higher heterozygosity and lower inbreeding, than sites in other regions. However, known small and likely extirpated populations in Michigan (e.g. DEL and BAS) have some of the highest levels of inbreeding, indicating that small population problems are ubiquitous across the range, and that inbreeding depression could be a cause of population extirpation.

We find evidence of both historical and recent inbreeding in eastern massasaugas. The majority of runs of homozygosity found in this study were small, indicating historical inbreeding and that populations throughout the range have maintained a small  $N_e$  even before widescale anthropogenic land conversion. This supports the results of Sovic et al. (2019), who found similar patterns in eastern massasauga populations outside of Michigan. All populations show some proportion of long runs of homozygosity, indicative of recent inbreeding over the past 200 years.

While reconstruction of historical  $N_e$  with GONE was limited to two mainland Michigan populations, these analyses show a historical decline before 200 years ago, followed by relatively stable population size, and recent declines associated with anthropogenic disturbance. GONE is less accurate in the distant past, so additional demographic inference using coalescence times could further explore the possibility and timing of declines in eastern massasauga populations before anthropogenic land use change. Additionally, we interpret the exact timing of events from GONE analyses with caution, as we are using a recombination rate approximation based on *Crotalus* rattlesnakes (Schield et al. 2020).

Snakes on Bois Blanc Island have elevated proportions of long ROH compared to other

populations, indicating inbreeding within the past 200 years. GONE analyses showed that the Bois Blanc population has been stable historically but has declined over the past 50 years. This is generally reflective of the Bois Blanc population having a distinct demographic trajectory compared to mainland sites.

The timing of inbreeding and demographic decline in our ROH and GONE analyses is based on our use of a generation time of three years. Generation time in eastern massasaugas could vary with latitude, with late-maturing population at higher latitudes. However, we choose to use a set generation time of three years, as it is unclear what populations should be considered late maturing (Faust et al., 2011). This could introduce biases in our data in that it is possible we are not comparing the same time bins across populations, so all dated events should be interpreted with caution.

Recent work has established that inbreeding depression does occur in eastern massasaugas (Clark et al. in prep). The two sites involved in that study, PCC and ELF are also represented in this study. We find that they have higher genetic diversity and lower  $F_{ROH}$  compared to other populations. Here, our estimates of  $F_{ROH}$  encapsulate both inbreeding due to non-random mating, and increases in homozygosity due to genetic drift, and are higher than inbreeding estimates from Clark et al., which used a different inbreeding estimator that primarily measured inbreeding due to non-random mating. Inbreeding depression in these two populations is likely driven by dispersal limitation resulting in mating within family groups. It is possible that this occurs in other populations of eastern massasaugas as well, and inbreeding depression could be further exacerbated by the effects of genetic drift.

#### *Impacts of land use conversion on eastern massasaugas*

Here, we document the impact of habitat loss due to land cover conversion on inbreeding

in eastern massasaugas. We found a significant relationship between inbreeding and anthropogenic disturbance within a 2x2 km area around each focal site. We interpret this as evidence that habitat conversion has reduced metapopulation dynamics and population sizes and resulted in increased inbreeding. This is also supported by our demographic inference results for the two mainland populations, which show evidence of decline close to the timing of local land cover conversion. It is also possible, though unlikely, that more inbred populations of eastern massasaugas are found in habitat more likely to be converted to another land use type.

We likely underestimate the impact of land cover on eastern massasauga populations. For example, we are using rough land use classifications: in reality, there are many types of natural habitats, and not all of them are suitable for eastern massasaugas. For example, encroachment of woody vegetation and invasive species like *Phragmites* in wetlands is a major threat to eastern massasauga populations (Dovčiak et al., 2013; Steen et al., 2015) that the resolution of our land cover data types does not account for.

### *Conclusions*

Eastern massasaugas are a unique Great Lakes species that face an uncertain future. They are threatened by climate change (Pomara et al., 2014), snake fungal disease (Allender et al., 2016; Hileman, Allender, et al., 2018; Tetzlaff et al., 2017), fire suppression and resulting ecological succession (Hileman, King, et al., 2018), habitat loss and fragmentation (Martin et al., 2021), human persecution (Faust et al., 2011) and inbreeding depression (Clark et al. *in prep*). We document that a substantial amount of genetic diversity is still found in the range core of eastern massasaugas. While these Michigan populations show signs of recent decline and inbreeding, they also may represent an important source of genetic diversity for species survival.

## **Acknowledgements**

This work was made possible through the efforts of many people who have participated in eastern massasauga population monitoring and sampling across the Great Lakes region, especially Jen Moore, members of the Moore Lab, Eric Hileman, Lisa Faust, and the Eastern Massasauga Rattlesnake Species Survival Plan working group. We are grateful to H. Lisle Gibbs and lab for early access to the eastern massasauga genome and fruitful discussion on eastern massasauga genomics. We would like to Tyler Linderoth and Rachel Toczydlowski for feedback on bioinformatics. This work was supported by Great Lakes Fish and Wildlife Restoration Act Award to SWF and JM, a Michigan State University Dissertation Completion Fellowship to MIC, and National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM137919 awarded to GSB. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Tables

**Table 3.1.** Sampling site abbreviation, geographic location, number of samples, and data source for whole genome sequences from eastern massasaugas rattlesnakes.

Site Code	State/Province, Country	No. samples	Data source
BBI	Michigan, USA	23	This study
ATL	Michigan, USA	2	This study
GRA	Michigan, USA	23	This study
HUR	Michigan, USA	8	This study
MAN	Michigan, USA	16	This study
SEV	Michigan, USA	12	This study
BAS	Michigan, USA	1	This study
GRF	Michigan, USA	9	This study
DEL	Michigan, USA	2	This study
HAL	Michigan, USA	13	This study
OTI	Michigan, USA	6	This study
PCC	Michigan, USA	25	This study
GOU	Michigan, USA	16	This study
ELF	Michigan, USA	25	This study
THR	Michigan, USA	10	This study
ONS	Michigan, USA	8	This study
SSSP	Illinois, USA	2	Mathur et al. 2023
SPVY	Ohio, USA	3	Mathur et al. 2023
PRDF	Ohio, USA	3	Mathur et al. 2023
CEBO	Ohio, USA	1	Mathur et al. 2023
KLDR	Ohio, USA	2	Mathur et al. 2023
WLRD	Ohio, USA	3	Mathur et al. 2023
ROME	Ohio, USA	1	Mathur et al. 2023
JENN	Pennsylvania, USA	4	Mathur et al. 2023
CCRO	New York, USA	2	Mathur et al. 2023
BPNP	Ontario, CA	1	Mathur et al. 2023

**Table 3.2.** Candidate set of multiple linear regressions.

	$\Delta AIC_c$	$\omega_i$	k	$-2(\log)\mathcal{L}$
<hr/>				
Full dataset:				
$F_{ROH} \sim \text{disturbance} + PC1 + PC2 + PC3 + PC4 + PC5 + PC6$	0	0.95	9	242.47
$F_{ROH} \sim PC1 + PC2 + PC3 + PC4 + PC5 + PC6$	5.82	0.05	8	238.47
<hr/>				
Disturbance > 0				
$F_{ROH} \sim \text{disturbance} + PC1 + PC2 + PC3 + PC4 + PC5 + PC6$	0	1	9	141.93
$F_{ROH} \sim PC1 + PC2 + PC3 + PC4 + PC5 + PC6$	17.03	0	8	132.25

## CHAPTER FOUR:

Inbreeding reduces fitness in spatially structured populations of a threatened rattlesnake

### **Abstract**

Small and fragmented populations are at high risk of extirpation, in part because of elevated inbreeding and subsequent inbreeding depression. A major conservation priority is to identify the mechanisms and extent of inbreeding depression in small populations. The eastern massasauga rattlesnake (*Sistrurus catenatus*) is federally listed as threatened, having experienced significant habitat fragmentation over the past 200 years. Here, we leverage long-term monitoring of two wild populations of eastern massasaugas in Michigan to estimate the extent of inbreeding in each population, identify mechanisms that generate inbreeding, and test for the impact of inbreeding on fitness. Using targeted genomic data and capture locations from over 1000 individuals, we find evidence of inbreeding and link inbreeding to spatial kinship structure within populations. We reconstruct multi-generational pedigrees for each population to measure reproductive output and use long term capture-recapture data to estimate individual apparent survival (i.e., the two major components of fitness). We find evidence of inbreeding depression in both fitness metrics. By combining genomics and long-term monitoring data, we are equipped answer fundamental questions about the nature of inbreeding depression in wild animal populations and provide important conservation context for future management of eastern massasaugas populations.

### **Introduction**

Anthropogenic land use changes have dramatically altered the size and distribution of natural populations (Haddad et al., 2015). In many species, contiguous ranges have been divided and populations are restricted to isolated habitat fragments. Small and isolated populations are

inherently more vulnerable to extinction than large, connected ones due to demographic and environmental stochasticity, reduced genetic variation and thus adaptive potential, and inbreeding depression, the reduction in fitness due to inbreeding (Gilpin & Soulé, 1986; Keller & Waller, 2002; Willi et al., 2006). Inbreeding depression can ultimately reduce population viability (Keller & Waller, 2002). Therefore, it is essential for the conservation of endangered species to identify how and why inbreeding depression occurs in small populations, and by how much fitness is reduced (Hedrick & Kalinowski, 2000; Lande, 1988; Mills & Smouse, 1994).

Two processes, genetic drift and inbreeding, generate inbreeding depression (Hedrick & Kalinowski, 2000; Keller & Waller, 2002). Genetic drift can cause the random loss or fixation of alleles within a population. Because the strength of drift is inversely correlated with effective population size, small populations are at greater risk of losing adaptive alleles and fixing deleterious alleles compared to larger populations, resulting in decreased genetic variation and population fitness (Lynch et al., 1995; Rich et al., 1979; Willi et al., 2006). Population-level inbreeding, or non-random mating between related individuals, reduces individual fitness by increasing the expression of deleterious recessive alleles in homozygous genotypes and/or reducing heterozygosity at loci that convey a fitness advantage in heterozygous genotypes (overdominance) (D. Charlesworth & Willis, 2009). Inbreeding can be generated by a variety of behaviors and life history traits, including mate choice (Thünken et al., 2007), the age distribution of reproducing individuals (Giesel, 1971), and limited dispersal distances (Ishihama et al., 2005).

Despite the experimental ubiquity and importance of understanding and quantifying inbreeding depression, it remains difficult to detect in natural populations, especially in those not easily monitored due to low encounter probabilities (Grueber et al., 2011; Hoy, Brzeski, et al.,

2023; Keller & Waller, 2002; Lavanchy et al., 2024). Inbreeding depression in wild populations is typically assessed by testing for a relationship between individual inbreeding, which approximates the autozygous proportion of the genome (i.e., identical-by-descent to itself) due to both inbreeding and genetic drift, and an individual fitness component (or proxy) using regression models (Hoy, Brzeski, et al., 2023). Although genomic data has allowed for accurate estimates of individual inbreeding in natural populations, measuring fitness directly remains difficult for most species (Alif et al., 2022; Kardos et al., 2016; Toczydlowski & Waller, 2023). Inbreeding depression varies in its impact on different components of fitness (Charpentier et al., 2007; Huisman et al., 2016), and commonly-used fitness proxies, like body condition, are often not strongly correlated with fitness (Wilder et al., 2016), leading to high false negative rates in tests for inbreeding depression in the wild. Here, we demonstrate the value of long-term population monitoring by testing for inbreeding depression using estimates of survival and reproductive success in wild populations of the elusive eastern massasauga (*Sistrurus catenatus*).

The eastern massasauga is a wetland-specialist rattlesnake that has experienced significant habitat fragmentation over the past 200 years (Szymanski et al., 2016). It is federally listed as threatened in the United States and as threatened or endangered under the Canadian Species at Risk Act. Although there is evidence that eastern massasaugas historically existed in small populations due to their association with naturally patchy wetland habitats, populations have experienced recent declines and extirpations associated with human land-use changes (Ochoa & Gibbs, 2021), and levels of natural connectivity have almost certainly been reduced. Contemporary populations of eastern massasaugas have small effective population sizes (Anthonysamy et al., 2022; Bradke, Hileman, et al., 2018; Sovic et al., 2019) and are vulnerable to habitat fragmentation, with agriculture, roads and other human modified habitats known to act

as barriers to dispersal (S. A. Martin et al., 2021, 2023; Moore & Gillingham, 2006; Paterson et al., 2019). Thus far, studies examining inbreeding have found no evidence of inbreeding depression in eastern massasaugas (Chiucchi & Gibbs, 2010; Sovic et al., 2019), but these studies used body condition as a fitness proxy, as measuring fitness in this species is difficult without extensive long-term monitoring.

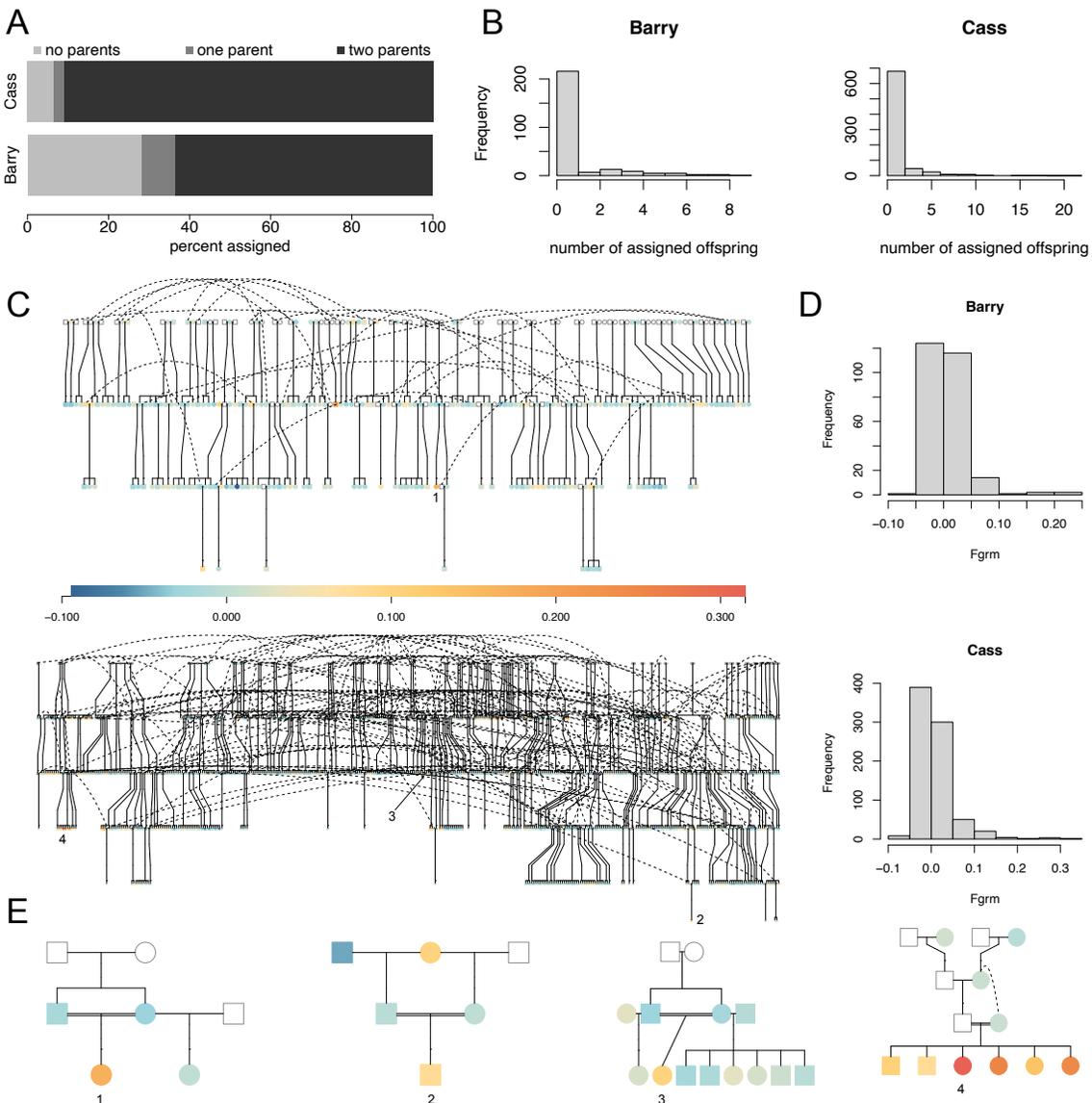
In this study, we merge over a decade of population monitoring with genetic pedigree reconstruction to quantify inbreeding in two populations of eastern massasaugas and test whether these small populations experience inbreeding depression. We (1) characterize the extent of individual inbreeding using genetic data, (2) analyze spatial structure that reinforces inbreeding, and (3) test for inbreeding depression using survival estimates from capture-recapture data and reproductive output measured from reconstructed wild pedigrees. We find that inbreeding impacts both survival and reproductive success in these populations, and that the geographic scale of dispersal and mating within populations may contribute to the prevalence of inbreeding. Our results highlight how aspects of life history, such as dispersal limitation, may contribute to inbreeding, and thereby its fitness consequences.

## **Results**

### *Genotyping individuals from >10 years of capture-recapture monitoring*

We focused on inbreeding dynamics in two populations of eastern massasaugas in southwestern Michigan, USA. We refer to these populations by their counties of origin: Barry and Cass. To estimate the extent of inbreeding and reconstruct pedigrees for each population, we used a RADseq + capture (i.e., RAPTURE) approach to target putatively neutral single nucleotide polymorphisms (SNPs) known to be polymorphic in these populations. Initial quality filtering resulted in a dataset of 5,607 SNPs in 1,056 individuals at an average depth of coverage

of 54X. After filtering for missing data and linkage, we retained a final dataset of 2,176 SNPs sequenced in 1,037 individuals across the two focal populations (Barry:  $n = 260$ , Cass:  $n = 776$ , approx. 50 km apart). Contemporary effective population sizes, estimated using linkage disequilibrium, were 48.07 (CI: 25.05–67.53) and 27.05 (CI: 21.87–31.63) at Barry and Cass Counties respectively.



**Figure 4.1.** (A) Figure shows the proportion of individuals in Barry and Cass Counties who were assigned no, one, or two parents in pedigree. (B) Histograms show the number of assigned offspring for Barry and Cass Counties. (C) Reconstructed pedigrees of eastern massasaugas in Barry County (top) and Cass County (bottom), Michigan, USA. Genotyped individuals are

**Figure 4.1. (cont'd)**

colored in accordance with  $F_{\text{grm}}$  and non-genotyped “dummy” individuals are outlined in gray. Pedigrees have been simplified for visualization, and some individuals appear multiple times. (D) Histograms showing the distribution of inbreeding in Barry and Cass Counties. (E) Family pedigrees show instances of close-kin mating that resulted in offspring with elevated  $F_{\text{grm}}$ . Numbers mark the location of focal individuals in both the family and population-wide pedigrees.

*Estimating reproductive output through pedigree reconstruction*

We reconstructed pedigrees separately for each population with the R package Sequoia (Huisman, 2017), using SNPs with minor allele frequencies above 0.1 in the focal population, as loci with more common minor alleles provide greater information for pedigree reconstruction (Barry: 1,452 SNPs, Cass: 1,472). We used a genotyping error rate of 0.0024 based on base-pair differences between 30 individuals sequenced twice. Most individuals had parents assigned in the reconstructed pedigrees (Figure 4.1, Barry County: 63.5% of 138 females and 119 males had both parents assigned, 8.5% had one parent assigned, 28.1% had no parents assigned, Cass County: 91.0% of 346 females and 278 males had both parents assigned, 2.6% had one parent assigned, 6.4% had no parents assigned). The discrepancy in parents assigned between populations reflects differences in population size and survey effort. Parents were either genotyped individuals or “dummy” individuals, the latter of which are inferred to exist via pedigree relationships between relatives. Dams and sires were assigned between one and nine pedigree offspring at Barry County and between one and 22 offspring at Cass County (Figure 4.1). A total of 65.9% ( $n = 91$ ) and 77.6% ( $n = 342$ ) of females and 75.6% ( $n = 90$ ) and 79.0% ( $n = 264$ ) of males at Barry and Cass Counties, respectively, did not have offspring assigned in the pedigree, suggesting substantial variation in reproductive success. Pedigrees encompassed a maximum of three generations at Barry County and five generations at Cass County.

Multiple lines of evidence suggest high confidence in inferred pedigree relationships. First, we included 265 snakes born in the lab to known dams (and subsequently released)

(Hileman et al., 2015; Stedman et al., 2016) and all of these relationships were correctly inferred in our pedigree reconstruction. Second, we only allowed types of relationships (e.g., both individuals genotyped, focal individual genotyped, parent individual dummy, etc.) for which the assignment confidence probability for the pairing was  $\geq 95\%$  (Tables S4.1 and S4. 2).

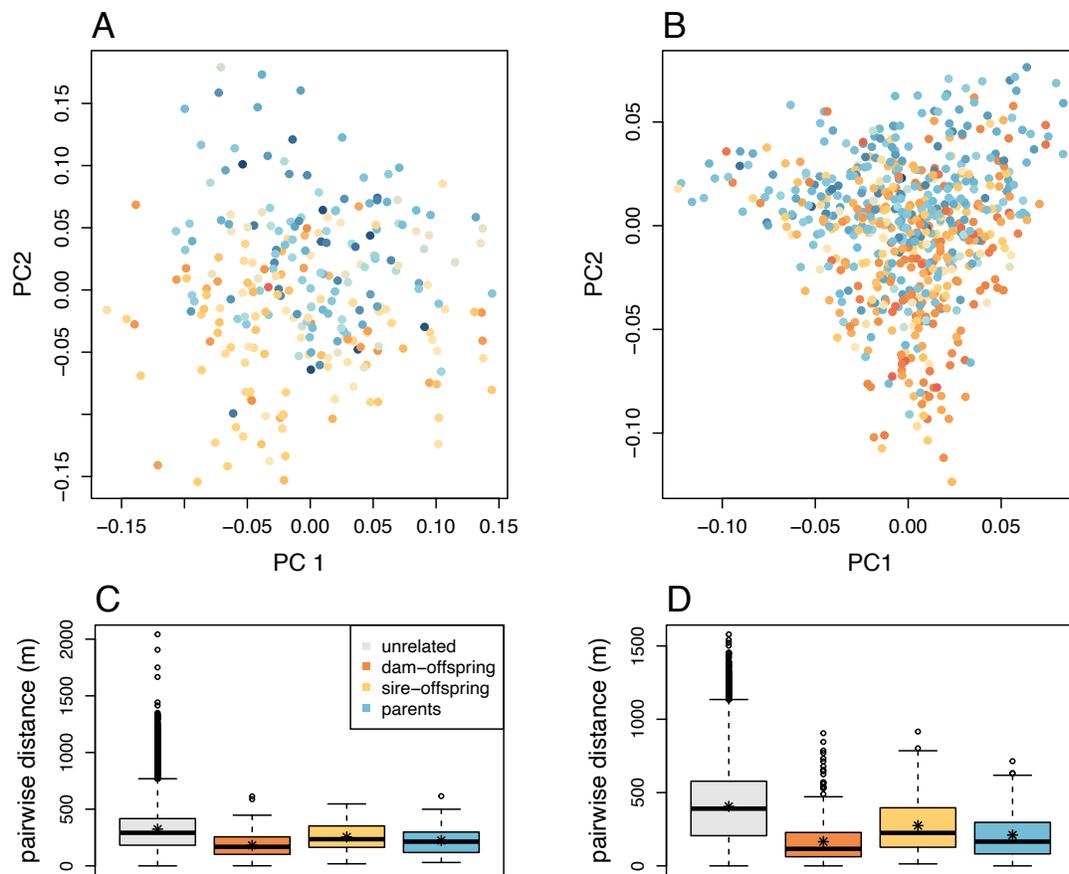
#### *Elevated genomic and pedigree inbreeding in eastern massasauga populations*

We estimated genomic inbreeding for individuals in each population separately using  $F_{\text{grm}}$ , a measure of individual inbreeding (Kardos et al., 2016, p. 20). Many calculations of genomic inbreeding exist; here we use  $F_{\text{grm}}$  because it places more weight on minor allele homozygotes that are less likely to be observed by chance compared to major allele homozygotes and has been used to detect inbreeding depression in other studies (Béréanos et al., 2016; Yang et al., 2011). Values of  $F_{\text{grm}}$  ranged from -0.0946 to 0.248 at Barry County and from -0.0724 to 0.315 at Cass County (Figure 4.1D). Negative values represent an excess of heterozygosity, reflecting individuals that are more outbred than expected given population allele frequencies. At Cass and Barry Counties, 10.2% and 7.3% of individuals had  $F_{\text{grm}}$  values greater than 0.05, respectively (Figure 4.1D). One individual from Cass County had an extremely elevated  $F_{\text{grm}}$  value (0.88), reflecting an excess of low frequency minor allele homozygous genotypes. We removed this individual from downstream analyses as a likely migrant.

We also used the inferred pedigree to calculate pedigree inbreeding coefficients (measures of the relatedness of inferred parents in the pedigree). Pedigree inbreeding ranged from 0 to 0.25 in both populations. At Barry County, one out of 37 unique pairs of genotyped parents were full siblings. At Cass County, seven of 100 unique pairs of genotyped parents half- or full-siblings. The offspring of the most related pairs of parents in the pedigree also had elevated  $F_{\text{grm}}$  ( $r(1034) = 0.466, p < 0.001$ ), Figure S4.1).

### *Spatial genetic structure detected within each population*

As expected, the Barry and Cass County sites were genetically differentiated from each other ( $F_{ST}=0.057$ ). In a principal component analysis, the first axis of variation (which explained 14.3% of the variation in the data) divided the counties (Figure S4.2). The potential migrant was an outlier in the PCA. Visualizing the spatial distribution of PCs from principal component analyses conducted separately for each site reveals fine-scale population structure within each site (Figure 4.2).



**Figure 4.2.** Top plots show principal component analyses for eastern massasaugas at Barry (A) and Cass (B) County sites. Points are colored according to an individual's centroid latitude at Barry or longitude at Cass, visualizing the main spatial axis of genetic variation within each site. Bottom plots show distributions of pairwise distances in meters between unrelated pairs without offspring (gray), dam-offspring pairs (orange), sire-offspring pairs (yellow), and parents with offspring in the pedigree (blue) for Barry (C) and Cass (D) County sites. The bar and star on each boxplot represent the median and mean of pairwise distances, respectively. All comparisons between dam-offspring, sire-offspring, and parent pairwise distance distributions and unrelated

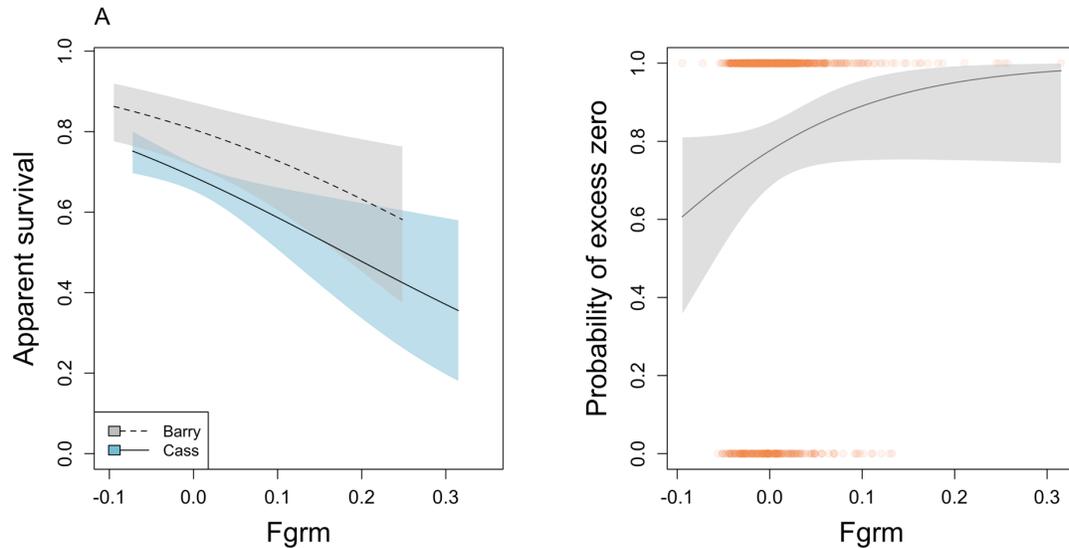
**Figure 4.2. (cont'd)**

distance distributions are significant ( $p < 0.05$ ) using Kolmogorov-Smirnov tests with Bonferroni corrected  $p$ -values.

To determine if this spatial pattern was the result of kinship clustering, we compared pairwise distances between related and unrelated individuals. Individuals recaptured in multiple locations were assigned a single location that was the centroid of all capture locations over time. Parents and offspring, as well as pairs of parents, were more likely to be closer together compared to pairs of unrelated individuals with no offspring (Figure 4.2, Bonferroni-corrected  $p$  for all tests  $< 0.005$ ). This mirrors a pattern of isolation by distance (Malécot, 1948; Wright, 1943) seen in the genetic data, where more genetically similar individuals are, on average, found closer together in geographic space (Figure S4.4).

*Inbreeding explains reduced individual survival*

We evaluated the effect of  $F_{\text{grm}}$  on apparent survival by building a set of candidate Cormack-Jolly-Seber (CJS) models that included 1,591 captures (Cass:  $n = 1,188$ ; Barry  $n = 403$ ) of 1,001 unique individuals (Table 3.1, Cass:  $n = 754$ ; Barry  $n = 247$ ) and determined the top-ranked model using  $AIC_c$  comparison. The top-ranked CJS model included negative additive effects of site (Cass) and  $F_{\text{grm}}$  and a positive additive effect of snout-vent length (SVL) to explain annual apparent survival (Table 3.1, Figures S3.4, S3.5). Barry County individuals had higher survival than Cass County individuals. To assess the effects of inbreeding on annual survival, we controlled for body size by setting SVL to 52 cm, which is the SVL of moderately-sized adult males and females at both sites. Cass County adult annual survival decreased from 0.75 to 0.36 as  $F_{\text{grm}}$  increased from -0.07 to 0.32 (Figure 4.3). Barry County adult annual survival decreased from 0.86 to 0.58 as  $F_{\text{grm}}$  increased from -0.09 to 0.25 (Figure 4.3).



**Figure 4.3.** (A) Effects of  $F_{grm}$  on annual apparent survival from capture-recapture models of eastern massasaugas from Barry (dashed line) and Cass (solid line) counties. Shaded bands represent 95% confidence intervals. To facilitate sex and site comparisons, variation in snout-vent length (SVL) was controlled for by holding it at 52 cm (moderately-sized adults at both sites). (B) Effect of  $F_{grm}$  on the probability of excess zeros in number of pedigree offspring in the zero-inflated model. The solid line represents the predicted probability of offspring. The shaded band represents the 95% confidence interval. Orange points show individuals that did (zeros) and did not (ones) have offspring.

#### *Inbreeding associated with reduced reproductive output*

To evaluate the effect of  $F_{grm}$  on total reproductive output we used zero-inflated negative binomial models with 826 genotyped individuals (Table 3.1, Cass:  $n = 589$ , Barry:  $n = 237$ ). These models incorporate a conditional and zero-inflated component to account for an excess of true and false zeros in the number of pedigree offspring (T. G. Martin et al., 2005). We expect true zeros to occur because of reproductive skew and false zeros to represent individuals never captured during the study. The top-ranked model from  $AIC_c$  comparison included effects of site, principal components 2-6 of the genetic PCA, years an individual could have contributed to the pedigree, sex, and  $F_{grm}$  in both the conditional and zero inflated components of the model (Table 3.1). The lowest-ranked model was the model that did not include  $F_{grm}$  as a predictor.

Individuals with higher inbreeding were more likely to have zero pedigree offspring (i.e.,

excess zeros), whereas snakes that had more years with the potential to contribute to the pedigree were less likely to have zero offspring (Figure 4.3, Figure S4.6). In the zero-inflated component of the model, a 0.05 increase in  $F_{\text{grm}}$  is associated with a 52.7% decrease in the odds of having any offspring in the pedigree. Years contributing to the pedigree and PC6 had a positive effect, and site (Barry) had a negative effect on the total number of offspring produced (Figure S4.6, S3.7). Inbreeding also had a negative effect on number of offspring, but the 85% confidence intervals overlapped zero, indicating it is not an important model parameter (Arnold, 2010).

## **Discussion**

We documented patterns of inbreeding and inbreeding depression in two wild populations of the threatened eastern massasauga. While most individuals in our study were not inbred, some individuals exhibited elevated levels of inbreeding, and inbreeding was associated with lower survival and reproductive success. We found evidence of fine-scale population structure within each site, which is a potential mechanism generating non-random mating and inbreeding depression. Other studies on eastern massasaugas and other rattlesnakes have documented the accumulation of putatively deleterious variants that likely contribute to inbreeding depression (Mathur et al., 2023), physical abnormalities thought to be a consequence of inbreeding (R. W. Clark et al., 2011; Gautschi et al., 2002; Madsen et al., 1996), and decreased litter size associated with low genetic variation (Madsen et al., 1996). However, to our knowledge, this is the first direct evidence of inbreeding depression in rattlesnakes. Our results join a growing body of work that suggests that inbreeding depression in animals is widespread and especially concerning given contemporary levels of habitat loss and fragmentation (Keller & Waller, 2002).

### *Extent of inbreeding*

While most individuals genotyped as a part of this study have inbreeding values near zero, indicating little to no inbreeding, more than 7% of individuals at both sites have  $F_{\text{grm}}$  values above 0.05, a value known to cause inbreeding depression in other taxa (Chen et al., 2016). Due to our large sample sizes, we were able to directly observe instances of true inbreeding in the pedigree (Figure 4.1). For example, the closest related parents were full siblings. We observe higher genomic inbreeding than pedigree inbreeding, likely due to the limited timescale encompassed by the pedigree (Kardos et al., 2016).

### *Fine-scale spatial structure generates inbreeding through non-random mating*

We found fine-scale spatial population structure within both sampled populations of eastern massasaugas (Figure 4.2). Fine-scale spatial genetic structure and spatially non-random mating has been connected to inbreeding in *Antechinus* marsupials and mountain brushtail possum (Blyton et al., 2015) and a decrease in fitness in other species including outcrossing plants (Ishihama et al., 2005). In other populations of flowering plants, spatially random mating acts to prevent inbreeding despite spatial structure (Coltman et al., 2003).

There are several facets of massasauga life history that could generate spatial structure and inbreeding: eastern massasaugas have limited dispersal and high site fidelity to overwintering sites (Harvey & Weatherhead, 2006), and neonates return to the areas near where they are born to overwinter (Jellen & Kowalski, 2007). Additionally, male snakes often return to specific areas each mating season to look for mates (Rouse et al., 2011). Thus, even though we observed individual snakes having moved up to 960 m between capture locations in this study, spatially restricted mating and parturition could generate spatial genetic structure. Additionally, anthropogenic land use changes reduce dispersal between populations (DiLeo et al., 2013; S. A.

Martin et al., 2021, 2023), perhaps increasing spatial structure within suitable habitat.

Understanding the demographic and dispersal histories of these populations could help elucidate the role anthropogenic barriers have played in the development of spatial structure within populations. Eastern massasaugas could be an example of a species especially susceptible to inbreeding and inbreeding depression because of their life history traits.

#### *Fitness consequences of inbreeding*

We detected a significant relationship between inbreeding and both survival and reproductive success in eastern massasaugas. Inbreeding had a strong negative effect on survival for both sexes at both sites. Inbred snakes could suffer higher mortality from predation, could be more susceptible to diseases, like snake fungal disease, which is known to be present, but rare, at these sites (Hileman, Allender, et al., 2018) and/or could experience greater overwintering mortality compared to non-inbred individuals. It is also possible that viability selection against inbred neonates could filter inbred individuals out of the population before they are detected, causing us to underestimate the extent of inbreeding depression on survival (Paijmans et al., 2024).

We find that the odds of having no offspring in the pedigree increased with inbreeding. We expect individuals are assigned no offspring due to both uncaptured individuals (“false zeros”) or because of reproductive skew (“true zeros”). We account for differences in survival between individuals by using the number of years an individual could have contributed to the pedigree as a covariate; thus the relationship between inbreeding and reproductive success is not driven by differences in parent survival. It is possible that inbred snakes have fewer mating opportunities, smaller litter size, higher neonate mortality, higher rate of failed gestation, and/or reproduce less frequently.

While genetic architecture underlying inbreeding depression in these population is unknown, our results suggest the presence and expression of deleterious recessive alleles within these populations. Ochoa and Gibbs (2021) found that eastern massasaugas in small populations outside of Michigan have fewer deleterious variants than their outbred congener *S. tergeminus*, suggesting that small historical population sizes have led to purging of some deleterious recessive alleles in eastern massasaugas. However, our results suggest that remaining deleterious recessive alleles are sufficient to generate inbreeding depression. Larger populations of eastern massasaugas in Michigan could also have not been as effective at purging compared to smaller populations near their range edges. Alternatively, a relationship between inbreeding and fitness could be generated via heterosis, whereby outbred individuals with high survival drive the relationship between inbreeding and survival. However, we did not observe many individuals with negative  $F_{gm}$ , indicating that heterosis is unlikely to produce the observed association.

In this study, we are only able to quantify the extent and impact of inbreeding due to nonrandom mating on eastern massasauga fitness. We expect that given small population sizes and likely recent isolation, these populations have also lost genetic diversity due to drift, which could further reduce the fitness of all individuals in the population as well as limit adaptive potential. This would be an interesting avenue for future investigation.

#### *Conclusions and conservation impacts*

It is essential to understand how fragmented and small populations persist on the landscape and identify future threats to persistence. We document the first evidence of inbreeding depression in eastern massasaugas, and in rattlesnakes in general. Long term population monitoring is essential for understanding the threats and subsequently conserving imperiled populations. While previous research has found that these populations are stable

(Bradke, Bailey, et al., 2018; Hileman, King, et al., 2018), our results indicate that there is sufficient deleterious variation in the gene pool to negatively impact fitness, and that this phenomenon might be common in eastern massasaugas in structured and isolated populations due to their life history. This highlights the importance of maintaining or restoring connectivity between populations for long-term population health.

## **Materials and Methods**

### *Sample collection*

This study focuses on two populations of eastern massasauga rattlesnakes in southwestern Michigan, 86 km apart: Cass and Barry Counties. Detailed descriptions of our eastern massasauga sampling and processing methods are described in Hileman et al. (2018), Bradke et al. (2018), and in the supplemental methods.

### *RAPTURE genotyping*

We extracted DNA from all unique blood samples collected until 2019 at Cass County and 2021 at Barry County. Twelve BestRAD libraries were prepared using custom RAPTURE baits and sequenced, containing 1,056 individuals (30 individuals sequenced as part of multiple libraries to act as technical duplicates) (Ali et al., 2016). Full description of our RAPTURE baits and bioinformatic pipeline is in the supplemental methods. Briefly, we aligned sequencing data to the eastern massasauga reference genome (Mathur et al., 2023) using BWA mem v. 07.17 (H. Li, 2013). Alignments were filtered and sorted using Samtools v. 1.9 (H. Li et al., 2009). We used the Stacks ref pipeline to call SNPs and remove PCR duplicates. We filtered SNPs using bcftools v.1.9.64 (H. Li, 2011) to retain SNPs with more than seven reads in 90% of individuals and with genotype quality scores greater than 19 in 90% of individuals. We filtered out SNPs with excess heterozygosity ( $p < 0.05$ ) and filtered to retain one randomly selected bi-allelic SNP

per targeted genomic region. Using custom R code, we removed SNPs with more than 10% missing data, and removed individuals with greater than 20% missing genotype calls or that had uneven distributions of reference and alternate allele read counts indicative of sample contamination. For principal components analysis, we filtered out SNPs with a minor allele frequency below 0.05. We filtered out SNPs with a minor allele frequency less than 0.1 separately for each site to reconstruct pedigrees. We calculated effective population  $N_e$  for each population using the strataG package in R with individuals from the same estimated birth year (Cass: 2013, Barry: 2015), using linkage disequilibrium based on Pearson correlation approximation and calculated 95% confidence intervals using jackknife resampling (Archer et al., 2017; R. K. Waples et al., 2016). We used the R package hierfstat to calculate Weir and Cockerham's estimate of  $F_{ST}$  (Goudet, 2005; Weir & Cockerham, 1984).

### *Spatial population structure*

For individuals with multiple captures, we calculated the centroid of their capture locations to use in spatial analyses using the "st\_centroid" function from the "sf" package (v. 1.0-9) in R (Pebesma, 2018). If an individual was only captured once, that capture location was used in spatial analyses. To visualize genetic structure within and between sites, we used principal component analyses implemented in custom R scripts. To assess the relationship between genetic and geographic distance, we plotted pairwise  $\pi$  at polymorphic sites against pairwise geographic distance. We compared distance distributions between closely related individuals (dam-offspring, sire-offspring) and parents (pairs with offspring in the pedigree) with the distribution of distances between unrelated individuals without offspring using Kolmogorov-Smirnov tests. All pairwise geographic distances were calculated using the "st\_distance" function from the package "sf" v.1.0-9 in R (Pebesma, 2018).

### *Quantifying inbreeding and genetic diversity*

To estimate individual inbreeding, we calculated  $F_{grm}$  using custom code in R (Bérénos et al., 2016; Yang et al., 2011). For a given individual,  $F_{grm}$  is defined as

$$F_{grm} = \frac{1}{L} \sum_i^L \frac{x_i^2 - (1 + 2p_i)x_i + 2p_i^2}{2p_i(1 - p_i)}$$

Where  $p_i$  is the population allele frequency at locus  $i$  and  $x_i$  is the number of copies of the reference allele in the focal individual. Multilocus heterozygosity and  $F_{grm}$  are highly correlated (Huisman et al., 2016). We calculated pairwise  $\pi$  at polymorphic sites using custom code. To measure inbreeding directly from the pedigree, we calculated pairwise kinship coefficients using the “CalcRped” function in Sequoia and identified related individuals with shared offspring in the pedigree. We calculated pedigree-based inbreeding coefficients using the “calcInbreeding” function in the pedigree v1.4.2 R package.

### *Testing for inbreeding depression: survival*

We evaluated the effects of site, sex, SVL, and  $F_{grm}$  on apparent survival using Cormack-Jolly-Seber (CJS) models implemented in the Program MARK v10.1 (White and Burnham 1999). We accounted for missing surveys in Barry (2009 and 2010) and Cass counties (2020) by creating dummy occasions and inserting ‘dots’ (.) indicating the missing years for affected individuals capture histories. We evaluated CJS assumptions using U-CARE (Choquet et al., 2009), replacing the dots with zeros, and stratified our capture histories into four site- and sex-specific groups. We detected significant transience and trap dependence for Cass County females, indicating an excess of snakes never detected after their initial capture and a trap shy effect in recapture probabilities. We assumed the transience and trap dependence effects in Cass County females were caused by unmodeled heterogeneity in apparent survival and recapture

probabilities, which we accounted for in the CJS models described below.

We treated sex and site as factors to explain apparent survival and recapture probabilities and  $F_{\text{grm}}$  as a static continuous individual covariate to explain apparent survival. We accounted for the approximately biennial change in adult female recapture probabilities (Hileman, King, et al., 2018) using a function that assigned individuals a ‘1’ if they were captured as an adult (i.e.,  $\geq 45$  cm SVL) two years prior to the current year and a ‘0’ if they were not captured two years ago as an adult (AFpriorcap2). To incorporate SVL as an individual-based temporally varying covariate in our CJS models, we constructed ten candidate models using the Fabens (1965) formulation of the von Bertalanffy (1938) growth model (Table S4.3)(Fabens, 1965; von Bertalanffy, 1938). We used individual records of date and length at initial and last capture, separated by  $\geq 1$  winter, to estimate the maximum length ( $L_{\infty}$ ) and growth coefficient (K) using nonlinear least squares in R version 4.3.2. We used  $AIC_c$  (Akaike, 1973) adjusted for small sample size via the R package *AICcmodavg* (Mazerolle 2023) for model selection. Models including uninformative parameters or that were simply nested versions of more parameterized ranking models were considered non-competing and discarded prior to assessing model support (Burnham and Anderson 2002, Arnold 2010). We used the top-ranked Faben’s model to predict SVL for all individuals with missing SVL values (Henderson et al. 2021).

We standardized and centered each covariate by subtracting off its mean and dividing it by two standard deviations so that continuous and binary regression coefficients were approximately on the same scale (Gelman, 2008). We constructed ten candidate CJS models using additive and interactive effects of the factors and covariates (Table 4.1). For the global CJS model, the apparent survival submodel included interactions between site and sex and between SVL and  $F_{\text{grm}}$  to explain apparent survival. The recapture submodel was highly embellished (42

parameters) to account for site differences in detection and sampling effort and was included in all candidate models. It included interactions between site and year and between sex and year, and an additive adult female behavioral response to recapture, (i.e., the AFpriorcap2 function). All candidate models were nested within the global model. We used the previously described model selection criteria to determine the most supported model.

### *Pedigree reconstruction*

Pedigrees for each site were reconstructed using the R package Sequoia v. 2.9.0 (Huisman, 2017), which assembles pedigrees based on genotype data, and approximate age and sex information. For individuals with unknown birth years, we approximated birth year using an individual's age class. For example, if an individual was captured as an adult, we inferred that the most recent possible birth year for that individual was three years previous, as eastern massasaugas reach sexual maturity by age 3 (Hileman, King, et al., 2018). We refined these estimates by inferring individual birth years using SVL and von Bertalanffy growth models (see Appendix B Supplemental Methods, Figures S3.8, S3.9). We used the inferred birth year +/- two years to narrow down the minimum and maximum birth years. We calculated genotyping error rate by comparing genotypes between technical replicates, after which duplicates were removed from the dataset. We evaluated our pedigree construction by validating that known dam-offspring relationships at the Cass County site ( $n = 265$ ) were called correctly. Of the offspring born in the lab, 46% were only recorded at birth and were not assigned offspring in the pedigree. These individuals were removed from offspring counts in downstream analyses. We also estimated our confidence in the reconstruction using the "EstConf()" function in Sequoia.

We used the pedigrees to estimate individual reproductive success (total number of offspring), and corrected these numbers to account for offspring of gravid females that gave birth

in captivity as part of another study (Stedman et al., 2016) such that we only counted a captive-born offspring towards a mother's reproductive success if that individual was recaptured after parturition or had offspring/grand-offspring of its own in the pedigree, indicating that it survived to adulthood.

*Testing for inbreeding depression: reproductive success*

To test for a relationship between reproductive success and  $F_{\text{grm}}$ , we built zero-inflated linear mixed models with a negative binomial distribution using the glmmTMB package in R. We constructed 3 models to test the effect of  $F_{\text{grm}}$  on the number of pedigree offspring. Our full model included: site, sex,  $F_{\text{grm}}$ , the number of years an individual could have contributed offspring to the pedigree, an interaction term between sex and  $F_{\text{grm}}$ , and 5 principal components (PC2-PC6, PC1 captured the effect of site) on number of offspring (Table 4.1). The number of years an individual could have contributed to the pedigree accounts for varying reproductive opportunities between individuals due to variation in longevity and birth year (see Supplemental Methods). PC2 through PC6 were included to account for genetic relatedness between individuals (Lavanchy et al., 2024). The interaction term between sex and  $F_{\text{grm}}$  tests the hypothesis that inbreeding depression is more severe for the homogametic sex. We scaled all numeric predictors by dividing each predictor by its standard deviation.

We included a zero-inflation term in our model to account for an excess of true and false zeros in the number of pedigree offspring (T. G. Martin et al., 2005). True zeros occur because of reproductive skew, and false zero represent individuals never captured. We used the same model for the conditional and zero-inflated components as our *a priori* expectation is that the same set of factors could contribute to both reproductive success and offspring count. We determined the most supported model using  $AIC_c$  comparison.

## **Acknowledgements**

This work would not be possible without the efforts and dedication of many people who have participated in population monitoring at both the Barry and Cass County sites over the past decade, especially Jen Moore, members of the Moore Lab and the Eastern Massasauga Rattlesnake Species Survival Plan working group. Eric Hileman led the capture-recapture models presented here, and Danielle Bradke led efforts to approximate individual ages. We are grateful to the Edward Lowe Foundation and Pierce Cedar Creek Institute for their dedication to land management and research access and H. Lisle Gibbs and lab for early access to the eastern massasauga genome. Members of the Fitzpatrick, Bradburd, and Winger labs provided valuable feedback on early stages of this manuscript. We would like to thank Madison Miller and Isabela Borges for assistance in the lab, and Tyler Linderoth and Rachel Toczydlowski for feedback on the bioinformatics and analyses. This work was supported by Great Lakes Fish and Wildlife Restoration Act Award to SWF and JM, NSF DEB Award to SWF, and National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM137919 (awarded to GSB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This is W.K. Kellogg Biological Station Contribution No. 2389.

## Tables

**Table 4.1.** Survival models: Candidate set of Cormack-Jolly-Seber models constructed from capture-recapture data collected from eastern massasauga populations in Cass and Barry counties between 2009–2023. Reproductive output models: Candidate set of zero-inflated negative binomial models.

Survival Models	$\Delta AIC_c$	$\omega_i$	k	$-2(\log)\mathcal{L}$
$\phi$ (site+SVL+Fgrm) $p$ (site*year+sex*year+AFpriorcap2)	0.00	0.315	46	3173.86
$\phi$ (site+sex+SVL+Fgrm) $p$ (site*year+sex*year+AFpriorcap2)	0.24	0.280	47	3171.98
$\phi$ (site+sex+SVL*Fgrm) $p$ (site*year+sex*year+AFpriorcap2)	0.53	0.242	48	3170.14
$\phi$ (site*sex+SVL*Fgrm) $p$ (site*year+sex*year+AFpriorcap2)	2.22	0.104	49	3169.70
$\phi$ (sex+SVL+Fgrm) $p$ (site*year+sex*year+AFpriorcap2)	5.28	0.023	46	3179.14
$\phi$ (SVL+Fgrm) $p$ (site*year+sex*year+AFpriorcap2)	5.53	0.020	45	3181.52
$\phi$ (site+SVL) $p$ (site*year+sex*year+AFpriorcap2)	6.96	0.010	45	3182.95
$\phi$ (site+sex+SVL) $p$ (site*year+sex*year+AFpriorcap2)	7.82	0.006	46	3181.68
$\phi$ (SVL) $p$ (site*year+sex*year+AFpriorcap2)	11.96	0.001	44	3190.06
$\phi$ (sex+SVL) $p$ (site*year+sex*year+AFpriorcap2)	12.31	0.001	45	3188.30
Reproductive Output Models				
offspring~ site+Fgrm+sex + PC2+PC3+PC4+PC5+PC6+years	0.00	0.55	21	1490.5
offspring~ site+Fgrm*sex + PC2+PC3+PC4+PC5+PC6+years	0.92	0.36	23	1487.22
offspring~ site + sex +PC2+PC3+PC4+PC5+PC6+years	3.93	0.08	19	1498.6

## REFERENCES

- Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, *60*(2), 255–265. <https://doi.org/10.1093/biomet/60.2.255>
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- 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. <https://doi.org/10.1534/genetics.115.183665>
- Alif, Ž., Dunning, J., Chik, H. Y. J., Burke, T., & Schroeder, J. (2022). What is the best fitness measure in wild populations? A case study on the power of short-term fitness proxies to predict reproductive value. *PLoS ONE*, *17*(4), e0260905. <https://doi.org/10.1371/journal.pone.0260905>
- Allender, M. C., Hileman, E. T., Moore, J., & Tetzlaff, S. (2016). Detection of Ophidiomyces, the causative agent of snake fungal disease, in the eastern massasauga (*Sistrurus catenatus*) in Michigan, USA, 2014. *Journal of Wildlife Diseases*, *52*(3), 694–698. <https://doi.org/10.7589/2015-12-333>
- Antao, T., Pérez-Figueroa, A., & Luikart, G. (2011). Early detection of population declines: High power of genetic monitoring using effective population size estimators. *Evolutionary Applications*, *4*(1), 144–154. <https://doi.org/10.1111/j.1752-4571.2010.00150.x>
- Anthonyamy, W. J. B., Dreslik, M. J., Baker, S. J., Davis, M. A., Douglas, M. R., Douglas, M. E., & Phillips, C. A. (2022). Limited gene flow and pronounced population genetic structure of Eastern Massasauga (*Sistrurus catenatus*) in a Midwestern prairie remnant. *PLOS ONE*, *17*(3), e0265666. <https://doi.org/10.1371/journal.pone.0265666>
- Archer, F. I., Adams, P. E., & Schneiders, B. B. (2017). stratag: An r package for manipulating, summarizing and analysing population genetic data. *Molecular Ecology Resources*, *17*(1), 5–11. <https://doi.org/10.1111/1755-0998.12559>
- Arnold, T. W. (2010). Uninformative Parameters and Model Selection Using Akaike's Information Criterion. *The Journal of Wildlife Management*, *74*(6), 1175–1178. <https://doi.org/10.1111/j.1937-2817.2010.tb01236.x>
- Batthey, C. J., Ralph, P. L., & Kern, A. D. (2020). Space is the place: Effects of continuous spatial structure on analysis of population genetic data. *Genetics*, *215*(1), 193–214. <https://doi.org/10.1534/genetics.120.303143>
- Baumdicker, F., Bisschop, G., Goldstein, D., Gower, G., Ragsdale, A. P., Tsambos, G., Zhu, S., Eldon, B., Ellerman, E. C., Galloway, J. G., Gladstein, A. L., Gorjanc, G., Guo, B., Jeffery, B., Kretschumar, W. W., Lohse, K., Matschiner, M., Nelson, D., Pope, N. S., ...

- Kelleher, J. (2022). Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*, 220(3), iyab229. <https://doi.org/10.1093/genetics/iyab229>
- Bay, R. A., Harrigan, R. J., Le Underwood, V., Lisle Gibbs, H., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, 359, 83–86. <https://doi.org/10.1126/science.aan4380>
- Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics*, 49, 433–456. <https://doi.org/10.1146/annurev-ecolsys-110617-062431>
- Béréanos, C., Ellis, P. A., Pilkington, J. G., & Pemberton, J. M. (2016). Genomic analysis reveals depression due to both individual and maternal inbreeding in a free-living mammal population. *Molecular Ecology*, 25(13), 3152–3168. <https://doi.org/10.1111/mec.13681>
- Bergeron, L. A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M. F., Quintard, B., Hoffman, J. I., Li, Z., St. Leger, J., Shao, C., Stiller, J., Gilbert, M. T. P., Schierup, M. H., & Zhang, G. (2023). Evolution of the germline mutation rate across vertebrates. *Nature*, 615(7951), Article 7951. <https://doi.org/10.1038/s41586-023-05752-y>
- Bi, K., Linderoth, T., Singhal, S., Vanderpool, D., Patton, J. L., Nielsen, R., Moritz, C., & Good, J. M. (2019). Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLOS Genetics*, 15(5), e1008119. <https://doi.org/10.1371/journal.pgen.1008119>
- Blyton, M. D. J., Banks, S. C., & Peakall, R. (2015). The effect of sex-biased dispersal on opposite-sexed spatial genetic structure and inbreeding risk. *Molecular Ecology*, 24(8), 1681–1695. <https://doi.org/10.1111/mec.13149>
- Bradke, D. R., Altobelli, J. T., Russell, A. L., Jaeger, C. P., & Moore, J. A. (2021). Low bottleneck detection in long-lived species despite lost genetic diversity: A case study of tuatara and eastern massasauga rattlesnakes. *Journal of Heredity*, 112(4), 346–356. <https://doi.org/10.1093/jhered/esab025>
- Bradke, D. R., Bailey, R. L., Bartman, J. F., Campa, H., Hileman, E. T., Krueger, C., Kudla, N., Lee, Y. M., Thacker, A. J., & Moore, J. A. (2018). Sensitivity analysis using site-specific demographic parameters to guide research and management of threatened eastern massasaugas. *Copeia*, 106(4), 600–610. <https://doi.org/10.1643/OT-18-059>
- Bradke, D. R., Hileman, E. T., Bartman, J. F., Faust, L. J., King, R. B., Kudla, N., & Moore, J. A. (2018). Implications of small population size in a threatened pitviper species. *Journal of Herpetology*, 52(4), 387–397. <https://doi.org/10.1670/18-026>
- Campana, S. E. (1999). Chemistry and composition of fish otoliths: Pathways, mechanisms and applications. *Marine Ecology Progress Series*, 188, 263–297. <https://doi.org/10.3354/meps188263>

- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1), s13742-015-0047–0048. <https://doi.org/10.1186/s13742-015-0047-8>
- Chapman, D. G. (1954). The estimation of biological populations. *The Annals of Mathematical Statistics*, 25(1), 1–15.
- Charbonnel, E., Daguin-Thiébaud, C., Caradec, L., Moittié, E., Gilg, O., Gavrilov, M. V., Strøm, H., Mallory, M. L., Morrison, R. I. G., Gilchrist, H. G., Leblois, R., Roux, C., Yearsley, J. M., Yannic, G., & Broquet, T. (2022). Searching for genetic evidence of demographic decline in an arctic seabird: Beware of overlapping generations. *Heredity*, 128(5), 364–376. <https://doi.org/10.1038/s41437-022-00515-3>
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195–205. <https://doi.org/10.1038/nrg2526>
- Charlesworth, D., & Willis, J. H. (2009). The genetics of inbreeding depression. *Nature Reviews Genetics*, 10(11), 783–796. <https://doi.org/10.1038/nrg2664>
- Charpentier, M. J. E., Widdig, A., & Alberts, S. C. (2007). Inbreeding depression in non-human primates: A historical review of methods used and empirical data. *American Journal of Primatology*, 69(12), 1370–1386. <https://doi.org/10.1002/ajp.20445>
- Chen, N., Cosgrove, E. J., Bowman, R., Fitzpatrick, J. W., & Clark, A. G. (2016). Genomic consequences of population decline in the endangered Florida scrub-jay. *Current Biology*, 26(21), 2974–2979. <https://doi.org/10.1016/j.cub.2016.08.062>
- Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*, 2(2), e107. <https://doi.org/10.1002/imt2.107>
- Chiu, J. E., & Gibbs, H. L. (2010). Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. *Molecular Ecology*, 19(24), 5345–5358. <https://doi.org/10.1111/j.1365-294X.2010.04860.x>
- Choquet, R., Lebreton, J., Gimenez, O., Reboulet, A., & Pradel, R. (2009). U-CARE: Utilities for performing goodness of fit tests and manipulating CAPTURE–RECAPTURE data. *Ecography*, 32(6), 1071–1074. <https://doi.org/10.1111/j.1600-0587.2009.05968.x>
- Clark, R. D., Catalano, K. A., Fitz, K. S., Garcia, E., Jaynes, K. E., Reid, B. N., Sawkins, A., Snead, A. A., Whalen, J. C., & Pinsky, M. L. (2023). The practice and promise of temporal genomics for measuring evolutionary responses to global change. *Molecular Ecology Resources*, 1755-0998.13789. <https://doi.org/10.1111/1755-0998.13789>
- Clark, R. W., Marchand, M. N., Clifford, B. J., Stechert, R., & Stephens, S. (2011). Decline of an isolated timber rattlesnake (*Crotalus horridus*) population: Interactions between climate change, disease, and loss of genetic diversity. *Biological Conservation*, 144(2), 886–891. <https://doi.org/10.1016/j.biocon.2010.12.001>

- Clegg, M. T., & Allard, R. W. (1973). Viability versus Fecundity Selection in the Slender Wild Oat, *Avena barbata* L. *Science*, *181*(4100), 667–668.  
<https://doi.org/10.1126/science.181.4100.667>
- Coltman, D. W., Pilkington, J. G., & Pemberton, J. M. (2003). Fine-scale genetic structure in a free-living ungulate population. *Molecular Ecology*, *12*(3), 733–742.  
<https://doi.org/10.1046/j.1365-294X.2003.01762.x>
- Congdon, J. D., Dunham, A. E., & Van Loben Sels, R. C. (1993). Delayed sexual maturity and demographics of Blanding’s turtles (*Emydoidea blandingii*): Implications for conservation and management of long-lived organisms. *Conservation Biology*, *7*(4), 826–833. <https://doi.org/10.1046/j.1523-1739.1993.740826.x>
- Conner, J. K., & Hartl, D. L. (2004). *A primer of ecological genetics* (Vol. 425). Sinauer Associates Sunderland, MA.
- Cook, R. E. (1979). Patterns of juvenile mortality and recruitment in plants. In O. T. Solbrig, S. Jain, G. B. Johnson, & P. H. Raven (Eds.), *Topics in Plant Population Biology* (pp. 207–231). Macmillan Education UK. [https://doi.org/10.1007/978-1-349-04627-0\\_10](https://doi.org/10.1007/978-1-349-04627-0_10)
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158.  
<https://doi.org/10.1093/bioinformatics/btr330>
- Davy, C. M., & Murphy, R. W. (2014). Conservation genetics of the endangered spotted turtle (*Clemmys guttata*) illustrate the risks of “bottleneck tests.” *Canadian Journal of Zoology*, *92*(2), 149–162. <https://doi.org/10.1139/cjz-2013-0188>
- Debinski, D. M., & Holt, R. D. (2000). A survey and overview of habitat fragmentation experiments. *Conservation Biology*, *14*(2), 342–355. <https://doi.org/10.1046/j.1523-1739.2000.98081.x>
- DiLeo, M. F., Rouse, J. D., Dávila, J. A., & Loughheed, S. C. (2013). The influence of landscape on gene flow in the eastern massasauga rattlesnake (*Sistrurus c. catenatus*): Insight from computer simulations. *Molecular Ecology*, *22*(17), 4483–4498.  
<https://doi.org/10.1111/mec.12411>
- Dolan, T. E., Palkovacs, E. P., Rogers, T. L., & Munch, S. B. (2023). Age structure augments the predictive power of time series for fisheries and conservation. *Canadian Journal of Fisheries and Aquatic Sciences*, *80*(5), 795–807. <https://doi.org/10.1139/cjfas-2022-0219>
- Dovčiak, M., Osborne, P. A., Patrick, D. A., & Gibbs, J. P. (2013). Conservation potential of prescribed fire for maintaining habitats and populations of an endangered rattlesnake *Sistrurus c. catenatus*. *Endangered Species Research*, *22*(1), 51–60.  
<https://doi.org/10.3354/esr00529>

- Doyle, J. M., Willoughby, J. R., Bell, D. A., Bloom, P. H., Bragin, E. A., Fernandez, N. B., Katzner, T. E., Leonard, K., & DeWoody, J. A. (2019). Elevated heterozygosity in adults relative to juveniles provides evidence of viability selection on eagles and falcons. *Journal of Heredity*, *110*(6), 696–706. <https://doi.org/10.1093/jhered/esz048>
- Eaton, M. J., & Link, W. A. (2011). Estimating age from recapture data: Integrating incremental growth measures with ancillary data to infer age-at-length. *Ecological Applications*, *21*(7), 2487–2497. <https://doi.org/10.1890/10-0626.1>
- England, P. R., Luikart, G., & Waples, R. S. (2010). Early detection of population fragmentation using linkage disequilibrium estimation of effective population size. *Conservation Genetics*, *11*(6), 2425–2430. <https://doi.org/10.1007/s10592-010-0112-x>
- Exposito-Alonso, M., Booker, T. R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C. C., Lang, P. L. M., Leventhal, L., Nogues-Bravo, D., Pagowski, V., Ruffley, M., Spence, J. P., Toro Arana, S. E., Weiß, C. L., & Zess, E. (2022). Genetic diversity loss in the Anthropocene. *Science*, *377*(6613), 1431–1435. <https://doi.org/10.1126/science.abn5642>
- Fabens, A. J. (1965). Properties and fitting of the von Bertalanffy growth curve. *Growth*, *29*, 265–289.
- Fallon, S. J., McDougall, A. J., Espinoza, T., Roberts, D. T., Brooks, S., Kind, P. K., Kennard, M. J., Bond, N., Marshall, S. M., Schmidt, D., & Hughes, J. (2019). Age structure of the Australian lungfish (*Neoceratodus forsteri*). *PLOS ONE*, *14*(1), e0210168. <https://doi.org/10.1371/journal.pone.0210168>
- Faust, L., Szymanski, J., & Redmer, M. (2011). *Range-wide extinction risk modeling for the eastern massasauga rattlesnake (Sistrurus catenatus catenatus)* [Technical Report to the U.S. Fish and Wildlife Service]. <https://doi.org/10.13140/RG.2.1.4182.8248>
- Felsenstein, J. (1971). Inbreeding and variance effective numbers in populations with overlapping generations. *Genetics*, *68*(4), 581–597. <https://doi.org/10.1093/genetics/68.4.581>
- Fisher, R. A. (1923). XXI.—On the Dominance Ratio. *Proceedings of the Royal Society of Edinburgh*, *42*, 321–341. <https://doi.org/10.1017/S0370164600023993>
- Fitzpatrick, S. W., Mittan-Moreau, C., Miller, M., & Judson, J. M. (2023). Genetic rescue remains underused for aiding recovery of federally listed vertebrates in the United States. *Journal of Heredity*, esad002. <https://doi.org/10.1093/jhered/esad002>
- Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: A review. *Genetics Research*, *66*(2), 95–107. <https://doi.org/10.1017/S0016672300034455>
- Gargiulo, R., Budde, K. B., & Heuertz, M. (2024). *Mind the lag: Understanding delayed genetic erosion*. <https://ecoevorxiv.org/repository/view/6756/>

- Gauthier, J., Pajkovic, M., Neuenschwander, S., Kaila, L., Schmid, S., Orlando, L., & Alvarez, N. (2020). Museomics identifies genetic erosion in two butterfly species across the 20th century in Finland. *Molecular Ecology Resources*, 20(5), 1191–1205. <https://doi.org/10.1111/1755-0998.13167>
- Gautschi, B., Widmer, A., Joshi, J., & Koella, J. C. (2002). Increased frequency of scale anomalies and loss of genetic variation in serially bottlenecked populations of the dice snake, *Natrix tessellata*. *Conservation Genetics*, 3(3), 235–245. <https://doi.org/10.1023/A:1019924514465>
- Gazey, W. J., & Staley, M. J. (1986). Population estimation from mark-recapture experiments using a sequential Bayes algorithm. *Ecology*, 67(4), 941–951. <https://doi.org/10.2307/1939816>
- Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27(15), 2865–2873. <https://doi.org/10.1002/sim.3107>
- Giesel, J. T. (1971). The relations between population structure and rate of inbreeding. *Evolution*, 25(3), 491–496. <https://doi.org/10.2307/2407347>
- Gilpin, M. E., & Soulé, M. E. (1986). Minimum viable populations: Processes of species extinction. In M. E. Soulé (Ed.), *Conservation biology: The science of scarcity and diversity* (pp. 19–34). Sinauer Associates, Inc.
- Goudet, J. (2005). Hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Grueber, C. E., Waters, J. M., & Jamieson, I. G. (2011). The imprecision of heterozygosity-fitness correlations hinders the detection of inbreeding and inbreeding depression in a threatened species. *Molecular Ecology*, 20(1), 67–79. <https://doi.org/10.1111/j.1365-294X.2010.04930.x>
- Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T. E., Sexton, J. O., Austin, M. P., Collins, C. D., Cook, W. M., Damschen, E. I., Ewers, R. M., Foster, B. L., Jenkins, C. N., King, A. J., Laurance, W. F., Levey, D. J., Margules, C. R., ... Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, 1(2), 1–10. <https://doi.org/10.1126/sciadv.1500052>
- Hailer, F., Helander, B., Folkestad, A. O., Ganusevich, S. A., Garstad, S., Hauff, P., Koren, C., Nygård, T., Volke, V., Vilà, C., & Ellegren, H. (2006). Bottlenecked but long-lived: High genetic diversity retained in white-tailed eagles upon recovery from population decline. *Biology Letters*, 2(2), 316–319. <https://doi.org/10.1098/rsbl.2006.0453>
- Haller, B. C., Galloway, J., Kelleher, J., Messer, P. W., & Ralph, P. L. (2019). Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Molecular Ecology Resources*, 19(2), 552–566. <https://doi.org/10.1111/1755-0998.12968>

- Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright – Fisher model. *Molecular Biology and Evolution*, 36(3), 632–637. <https://doi.org/10.1093/molbev/msy228>
- Haller, B. C., & Messer, P. W. (2023). SLiM 4: Multispecies eco-evolutionary modeling. *The American Naturalist*, 201(5), E127–E139. <https://doi.org/10.1086/723601>
- Harvey, D. S., & Weatherhead, P. J. (2006). Hibernation site selection by eastern massasauga rattlesnakes (*Sistrurus catenatus catenatus*) near their northern range limit. *Journal of Herpetology*, 40(1), 66–73.
- Hausmann, M. F., & Vleck, C. M. (2002). Telomere length provides a new technique for aging animals. *Oecologia*, 130(3), 325–328. <https://doi.org/10.1007/s00442-001-0827-y>
- Hedrick, P. W., & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, 31(Volume 31, 2000), 139–162. <https://doi.org/10.1146/annurev.ecolsys.31.1.139>
- Heyrend, F. L., & Call, A. (1951). Growth and age in western striped racer and Great Basin rattlesnake. *Herpetologica*, 7(1), 28–40.
- Hileman, E. T., Allender, M. C., Bradke, D. R., Faust, L. J., Moore, J. A., Ravesi, M. J., & Tetzlaff, S. J. (2018). Estimation of *Ophidiomyces* prevalence to evaluate snake fungal disease risk. *The Journal of Wildlife Management*, 82(1), 173–181. <https://doi.org/10.1002/jwmg.21345>
- Hileman, E. T., Bradke, D. R., Delaney, D. M., & King, R. B. (2015). Protection by association: Implications of scent trailing in neonate eastern massasaugas (*Sistrurus catenatus*). *Herpetological Conservation and Biology*.
- Hileman, E. T., King, R. B., Adamski, J. M., Anton, T. G., Bailey, R. L., Baker, S. J., Bieser, N. D., Bell, T. A., Bissell, K. M., Bradke, D. R., Campa, H., Casper, G. S., Cedar, K., Cross, M. D., DeGregorio, B. A., Dreslik, M. J., Faust, L. J., Harvey, D. S., Hay, R. W., ... Yagi, A. (2017). Climatic and geographic predictors of life history variation in Eastern Massasauga (*Sistrurus catenatus*): A range-wide synthesis. *PLOS ONE*, 12(2), e0172011. <https://doi.org/10.1371/journal.pone.0172011>
- Hileman, E. T., King, R. B., & Faust, L. J. (2018). Eastern massasauga demography and extinction risk under prescribed-fire scenarios. *Journal of Wildlife Management*, 82(5), 977–990. <https://doi.org/10.1002/jwmg.21457>
- Hill, W. G. (1972). Effective size of populations with overlapping generations. *Theoretical Population Biology*, 3(3), 278–289. [https://doi.org/10.1016/0040-5809\(72\)90004-4](https://doi.org/10.1016/0040-5809(72)90004-4)
- Hill, W. G. (1979). A note on effective population size with overlapping generations. *Genetics*, 92(1), 317–322.

- Hinton, M. S., McMillan, B. R., Hersey, K. R., & Larsen, R. T. (2023). Estimating age of mule deer in the field: Can we move beyond broad age categories? *PLOS ONE*, *18*(7), e0284565. <https://doi.org/10.1371/journal.pone.0284565>
- Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479–485. <https://doi.org/10.1038/nature09670>
- Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62–82. <https://doi.org/10.1111/mec.15720>
- Hollenbeck, C. M., Portnoy, D. S., & Gold, J. R. (2016). A method for detecting recent changes in contemporary effective population size from linkage disequilibrium at linked and unlinked loci. *Heredity*, *117*(4), Article 4. <https://doi.org/10.1038/hdy.2016.30>
- Holmes, E. E., & York, A. E. (2003). Using age structure to detect impacts on threatened populations: A case study with Steller sea lions. *Conservation Biology*, *17*(6), 1794–1806. <https://doi.org/10.1111/j.1523-1739.2003.00191.x>
- Hoy, S. R., Brzeski, K. E., Vucetich, L. M., Peterson, R. O., & Vucetich, J. A. (2023). The difficulty of detecting inbreeding depression and its effect on conservation decisions. *Journal of Heredity*, esad080. <https://doi.org/10.1093/jhered/esad080>
- Hoy, S. R., Hedrick, P. W., Peterson, R. O., Vucetich, L. M., Brzeski, K. E., & Vucetich, J. A. (2023). The far-reaching effects of genetic process in a keystone predator species, grey wolves. *Science Advances*, *9*(34), eadc8724. <https://doi.org/10.1126/sciadv.adc8724>
- Hsu, J. L., Kam, S., Tammone, M. N., Lacey, E. A., & Hadly, E. A. (2017). Rapid increase in genetic diversity in an endemic Patagonian tuco-tuco following a recent volcanic eruption. *Journal of Mammalogy*, *98*(3), 779–792. <https://doi.org/10.1093/jmammal/gyx008>
- Huffmeyer, A. A., Sikich, J. A., Vickers, T. W., Riley, S. P. D., & Wayne, R. K. (2022). First reproductive signs of inbreeding depression in Southern California male mountain lions (*Puma concolor*). *Theriogenology*, *177*, 157–164. <https://doi.org/10.1016/j.theriogenology.2021.10.016>
- Huisman, J. (2017). Pedigree reconstruction from SNP data: Parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, *17*(5), 1009–1024. <https://doi.org/10.1111/1755-0998.12665>
- Huisman, J., Kruuk, L. E. B., Ellis, P. A., Clutton-Brock, T., & Pemberton, J. M. (2016). Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences*, *113*(13), 3585–3590. <https://doi.org/10.1073/pnas.1518046113>
- Ishihama, F., Ueno, S., Tsumura, Y., & Washitani, I. (2005). Gene flow and inbreeding depression inferred from fine-scale genetic structure in an endangered heterostylous

- perennial, *Primula sieboldii*. *Molecular Ecology*, *14*(4), 983–990.  
<https://doi.org/10.1111/j.1365-294X.2005.02483.x>
- Jellen, B. C., & Kowalski, M. J. (2007). Movement and growth of neonate eastern massasaugas (*Sistrurus catenatus*). *Copeia*, *2007*(4), 994–1000.
- Jensen, E. L., Edwards, D. L., Garrick, R. C., Miller, J. M., Gibbs, J. P., Cayot, L. J., Tapia, W., Caccone, A., & Russello, M. A. (2018). Population genomics through time provides insights into the consequences of decline and rapid demographic recovery through head-starting in a Galapagos giant tortoise. *Evolutionary Applications*, *11*(10), 1811–1821.  
<https://doi.org/10.1111/eva.12682>
- Jones, J. A. B., Lenart, A., & Baudisch, A. (2018). Complexity of the relationship between life expectancy and overlap of lifespans. *PLOS ONE*, *13*(7), e0197985.  
<https://doi.org/10.1371/journal.pone.0197985>
- Kamm, J., Terhorst, J., Durbin, R., & Song, Y. S. (2020). Efficiently inferring the demographic history of many populations with allele count data. *Journal of the American Statistical Association*, *115*(531), 1472–1487. <https://doi.org/10.1080/01621459.2019.1635482>
- 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. <https://doi.org/10.1111/eva.12414>
- Kelleher, J., Etheridge, A. M., & McVean, G. (2016). Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Computational Biology*, *12*(5), 1–22.  
<https://doi.org/10.1371/journal.pcbi.1004842>
- Kelleher, J., Thornton, K. R., Ashander, J., & Ralph, P. L. (2018). Efficient pedigree recording for fast population genetics simulation. *PLoS Computational Biology*, *14*(11), e1006581.  
<https://doi.org/10.1371/journal.pcbi.1006581>
- Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230–241. [https://doi.org/10.1016/S0169-5347\(02\)02489-8](https://doi.org/10.1016/S0169-5347(02)02489-8)
- Kettle, C. J., Hollingsworth, P. M., Jaffré, T., Moran, B., & Ennos, R. A. (2007). Identifying the early genetic consequences of habitat degradation in a highly threatened tropical conifer, *Araucaria nemorosa* Laubenfels. *Molecular Ecology*, *16*(17), 3581–3591.  
<https://doi.org/10.1111/j.1365-294X.2007.03419.x>
- Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, *21*(4), 1359–1368. <https://doi.org/10.1111/1755-0998.13326>
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>

- Kramer, A. M., Dennis, B., Liebhold, A. M., & Drake, J. M. (2009). The evidence for Allee effects. *Population Ecology*, *51*(3), 341–354. <https://doi.org/10.1007/s10144-009-0152-6>
- Kuo, C. H., & Janzen, F. J. (2003). BOTTLESIM: A bottleneck simulation program for long-lived species with overlapping generations. *Molecular Ecology Notes*, *3*(4), 669–673. <https://doi.org/10.1046/j.1471-8286.2003.00532.x>
- Labonne, J., Kaeuffer, R., Guéraud, F., Zhou, M., Manicki, A., & Hendry, A. P. (2016). From the bare minimum: Genetics and selection in populations founded by only a few parents. *Evolutionary Ecology Research*, *17*(1), 21–34.
- Lande, R. (1988). Genetics and demography in biological conservation. *Science*, *241*(4872), 1455–1460.
- Lavanchy, E., Weir, B. S., & Goudet, J. (2024). Detecting inbreeding depression in structured populations. *Proceedings of the National Academy of Sciences*, *121*(19), e2315780121. <https://doi.org/10.1073/pnas.2315780121>
- Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular Biology and Evolution*, *32*(10), 2798–2800. <https://doi.org/10.1093/molbev/msv150>
- Lewanski, A. L., Grundler, M. C., & Bradburd, G. S. (2024). The era of the ARG: An introduction to ancestral recombination graphs and their significance in empirical evolutionary genomics. *PLOS Genetics*, *20*(1), e1011110. <https://doi.org/10.1371/journal.pgen.1011110>
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Preprint*, *00*(00), 1–3.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, X., Tian, H., Lu, C., & Pan, S. (2023). Four-century history of land transformation by humans in the United States (1630–2020): Annual and 1 km grid data for the HIStory of LAND changes (HISLAND-US). *Earth System Science Data*, *15*(2), 1005–1035. <https://doi.org/10.5194/essd-15-1005-2023>
- Lippé, C., Dumont, P., & Bernatchez, L. (2006). High genetic diversity and no inbreeding in the endangered copper redhorse, *Moxostoma hubbsi* (Catostomidae, Pisces): The positive sides of a long generation time. *Molecular Ecology*, *15*(7), 1769–1780. <https://doi.org/10.1111/j.1365-294X.2006.02902.x>

- Lynch, M., Conery, J., & Burger, R. (1995). Mutation accumulation and the extinction of small populations. *The American Naturalist*, 146(4), 489–518. <https://doi.org/10.1086/285812>
- Madsen, T., Stille, B., & Shine, R. (1996). Inbreeding depression in an isolated population of adders *Vipera berus*. *Biological Conservation*, 75(2), 113–118. [https://doi.org/10.1016/0006-3207\(95\)00067-4](https://doi.org/10.1016/0006-3207(95)00067-4)
- Malécot, G. (1948). *Les Mathématiques de L'hérité*. Masson and Cie.
- Martin, S. A., Lipps, G. J., & Gibbs, H. L. (2021). Pedigree-based assessment of recent population connectivity in a threatened rattlesnake. *Molecular Ecology Resources*, 21(6), 1820–1832. <https://doi.org/10.1111/1755-0998.13383>
- Martin, S. A., Peterman, W. E., Lipps Jr., G. J., & Gibbs, H. L. (2023). Inferring population connectivity in eastern massasauga rattlesnakes (*Sistrurus catenatus*) using landscape genetics. *Ecological Applications*, 33(2), e2793. <https://doi.org/10.1002/eap.2793>
- Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., & Possingham, H. P. (2005). Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. *Ecology Letters*, 8(11), 1235–1246. <https://doi.org/10.1111/j.1461-0248.2005.00826.x>
- Mathur, S., Mason, A. J., Bradburd, G. S., & Gibbs, H. L. (2023). Functional genomic diversity is correlated with neutral genomic diversity in populations of an endangered rattlesnake. *Proceedings of the National Academy of Sciences*, 120(43), e2303043120. <https://doi.org/10.1073/pnas.2303043120>
- McCallum, H. (2012). Disease and the dynamics of extinction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1604), 2828–2839. <https://doi.org/10.1098/rstb.2012.0224>
- McCann, M. T. (1979). The plant tension zone in Michigan [M.A., Western Michigan University]. In *ProQuest Dissertations and Theses* (302948491). ProQuest Dissertations & Theses Global; ProQuest Dissertations & Theses Global Closed Collection. <https://ezproxy.msu.edu/login?url=https://www.proquest.com/dissertations-theses/plant-tension-zone-michigan/docview/302948491/se-2?accountid=12598>
- McCarty, J. P. (2001). Ecological Consequences of Recent Climate Change. *Conservation Biology*, 15(2), 320–331. <https://doi.org/10.1046/j.1523-1739.2001.015002320.x>
- Miller, P. (2005). *Population viability assessment for the eastern massasauga rattlesnake (Sistrurus catenatus catenatus) on the Bruce Peninsula, Ontario, Canada*. Conservation Breeding Specialist Group.
- Mills, L. S., & Smouse, P. E. (1994). Demographic consequences of inbreeding in remnant populations. *The American Naturalist*, 144(3), 412–431.

- Moore, J. A., & Gillingham, J. C. (2006). Spatial ecology and multi-scale habitat selection by a threatened rattlesnake: The Eastern Massasauga (*Sistrurus catenatus catenatus*). *Copeia*, 2006(4), 742–751. [https://doi.org/10.1643/0045-8511\(2006\)6\[742:SEAMHS\]2.0.CO;2](https://doi.org/10.1643/0045-8511(2006)6[742:SEAMHS]2.0.CO;2)
- Nadachowska-Brzyska, K., Koneczal, M., & Babik, W. (2022). Navigating the temporal continuum of effective population size. *Methods in Ecology and Evolution*, 13(1), 22–41. <https://doi.org/10.1111/2041-210X.13740>
- Nakamura, S., Yamazaki, J., Matsumoto, N., Inoue-Murayama, M., Qi, H., Yamanaka, M., Nakanishi, M., Yanagawa, Y., Sashika, M., Tsubota, T., Ito, H., & Shimozuru, M. (2023). Age estimation based on blood DNA methylation levels in brown bears. *Molecular Ecology Resources*, 1755-0998.13788. <https://doi.org/10.1111/1755-0998.13788>
- Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C., & Durbin, R. (2016). BCFtools/RoH: A hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics*, 32(11), 1749–1751. <https://doi.org/10.1093/bioinformatics/btw044>
- Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), 5269–5273. <https://doi.org/10.1073/pnas.76.10.5269>
- Newman, D., & Pilson, D. (1997). Increased probability of extinction due to decreased genetic effective population size: Experimental populations of *Clarkia pulchella*. *Evolution*, 51(2), 354–362. <https://doi.org/10.1111/j.1558-5646.1997.tb02422.x>
- Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution*, 47(5), 1329–1341. <https://doi.org/10.1111/j.1558-5646.1993.tb02158.x>
- Ochoa, A., & Gibbs, H. L. (2021). Genomic signatures of inbreeding and mutation load in a threatened rattlesnake. *Molecular Ecology*, 30(21), 5454–5469. <https://doi.org/10.1111/mec.16147>
- Paijmans, A. J., Berthelsen, A. L., Nagel, R., Cristaller, F., Kröcker, N., Forcada, J., & Hoffman, J. I. (2024). Little evidence for inbreeding depression for birth mass, survival and growth in Antarctic fur seal pups (p. 2024.01.12.575355). bioRxiv. <https://doi.org/10.1101/2024.01.12.575355>
- Paterson, J. E., Baxter-Gilbert, J., Beaudry, F., Carstairs, S., Chow-Fraser, P., Edge, C. B., Lentini, A. M., Litzgus, J. D., Markle, C. E., McKeown, K., Moore, J. A., Refsnider, J. M., Riley, J. L., Rouse, J. D., Seburn, D. C., Zimmerling, J. R., & Davy, C. M. (2019). Road avoidance and its energetic consequences for reptiles. *Ecology and Evolution*, 9(17), 9794–9803. <https://doi.org/10.1002/ece3.5515>
- Pebesma, E. J. (2018). Simple features for R: standardized support for spatial vector data. *R J.*, 10(1), 439.

- Peery, Z. M., Kirby, R., Reid, B. N., Stoelting, R., Doucet-Béer, E., Robinson, S., Vásquez-Carrillo, C., Pauli, J. N., & Palsboll, P. J. (2012). Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, *21*(14), 3403–3418. <https://doi.org/10.1111/j.1365-294X.2012.05635.x>
- Pereira, F. B., Sebbenn, A. M., Boshier, D. H., Rossini, B. C., Marino, C. L., Freitas, M. L. M., Rosa, J. R. B. F., Vidal, E., & Tambarussi, E. V. (2023). Gene flow, mating patterns and inbreeding depression in *Roupala montana* var. *Brasiliensis*, a neotropical timber species. *New Forests*. <https://doi.org/10.1007/s11056-023-10009-7>
- Pham, C. H., Price, J. J., Tallant, J. M., & Karowe, D. N. (2022). Climate change is predicted to reduce sympatry among North American wood-warblers. *Ornithological Applications*, *124*(4), duac025.
- Pike, D. A., Pizzatto, L., Pike, B. A., & Shine, R. (2008). Estimating survival rates of uncatchable animals: The myth of high juvenile mortality in reptiles. *Ecology*, *89*(3), 607–611. <https://doi.org/10.1890/06-2162.1>
- Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187). <https://doi.org/10.1126/science.1246752>
- Pomara, L. Y., LeDee, O. E., Martin, K. J., & Zuckerberg, B. (2014). Demographic consequences of climate change and land cover help explain a history of extirpations and range contraction in a declining snake species. *Global Change Biology*, *20*(7), 2087–2099. <https://doi.org/10.1111/gcb.12510>
- Pyle, P. (1997). Molt Limits in North American. *North American Bird Bander*, *22*(2), 49–89.
- Ralston, J., & Kirchman, J. J. (2013). Predicted range shifts in North American boreal forest birds and the effect of climate change on genetic diversity in blackpoll warblers (*Setophaga striata*). *Conservation Genetics*, *14*(2), 543–555. <https://doi.org/10.1007/s10592-012-0418-y>
- Reid, B. N., & Pinsky, M. L. (2022). Simulation-based evaluation of methods, data types, and temporal sampling schemes for detecting recent population declines. *Integrative and Comparative Biology*, icac144. <https://doi.org/10.1093/icb/icac144>
- Rich, S. S., Bell, A. E., & Wilson, S. P. (1979). Genetic drift in small populations of *Tribolium*. *Evolution*, *33*(2), 579–584. <https://doi.org/10.2307/2407781>
- Robinson, J. A., Räikkönen, J., Vucetich, L. M., Vucetich, J. A., Peterson, R. O., Lohmueller, K. E., & Wayne, R. K. (2019). Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Science Advances*, *5*(5), eaau0757. <https://doi.org/10.1126/sciadv.aau0757>

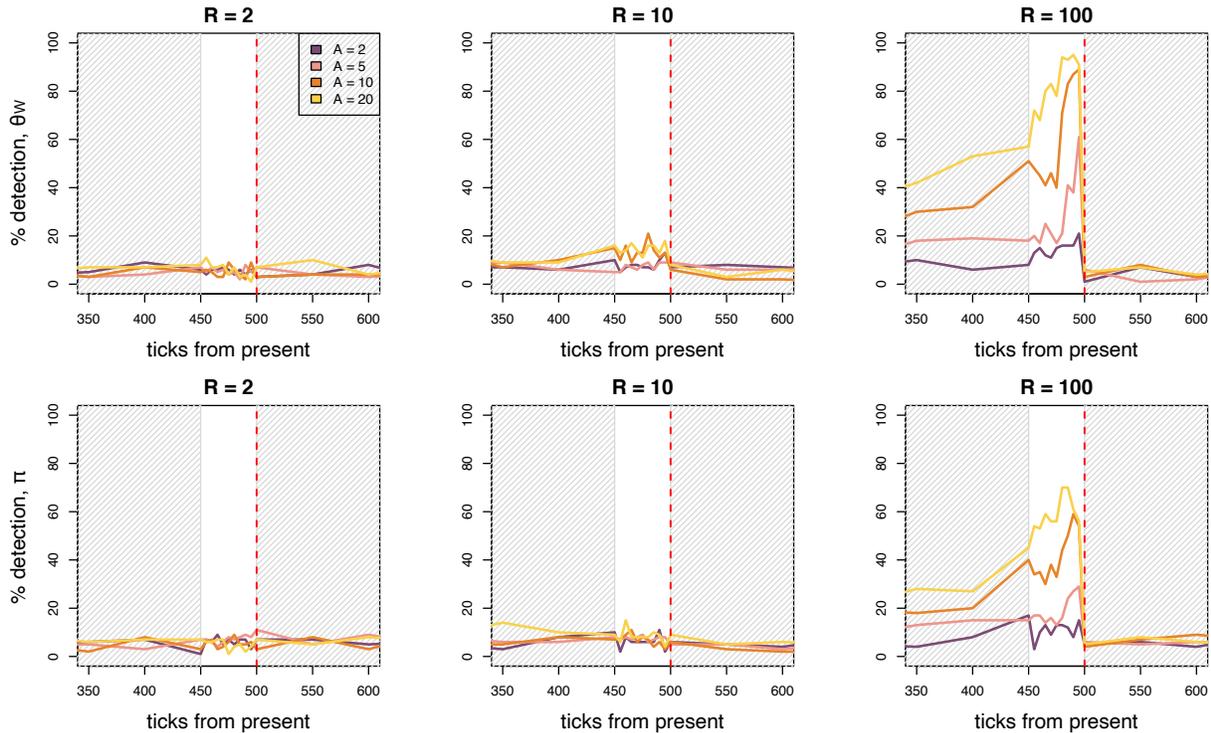
- Rouse, J. D., Willson, R. J., Black, R., & Brooks, R. J. (2011). Movement and spatial dispersion of *Sistrurus catenatus* and *Heterodon platirhinos*: Implications for interactions with roads. *Copeia*, 2011(3), 443–456.
- Santiago, E., Novo, I., Pardiñas, A. F., Saura, M., Wang, J., & Caballero, A. (2020). Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Molecular Biology and Evolution*, 37(12), 3642–3653. <https://doi.org/10.1093/molbev/msaa169>
- Saremi, N. F., Supple, M. A., Byrne, A., Cahill, J. A., Coutinho, L. L., Dalén, L., Figueiró, H. V., Johnson, W. E., Milne, H. J., O'Brien, S. J., O'Connell, B., Onorato, D. P., Riley, S. P. D., Sikich, J. A., Stahler, D. R., Vilella, P. M. S., Vollmers, C., Wayne, R. K., Eizirik, E., ... Shapiro, B. (2019). Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, 10(1), 4769. <https://doi.org/10.1038/s41467-019-12741-1>
- Schild, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake recombination landscapes are concentrated in functional regions despite PRDM9. *Molecular Biology and Evolution*, 37(5), 1272–1294. <https://doi.org/10.1093/molbev/msaa003>
- Schmidt, D. J., Fallon, S., Roberts, D. T., Espinoza, T., McDougall, A., Brooks, S. G., Kind, P. K., Bond, N. R., Kennard, M. J., & Hughes, J. M. (2018). Monitoring age-related trends in genomic diversity of Australian lungfish. *Molecular Ecology*, 27(16), 3231–3241. <https://doi.org/10.1111/mec.14791>
- Segelbacher, G., Strand, T. M., Quintela, M., Axelsson, T., Jansman, H. A. H., Koelewijn, H.-P., & Höglund, J. (2014). Analyses of historical and current populations of black grouse in Central Europe reveal strong effects of genetic drift and loss of genetic diversity. *Conservation Genetics*, 15(5), 1183–1195. <https://doi.org/10.1007/s10592-014-0610-3>
- Shroder, J. F. (1980). Dendrogeomorphology: Review and new techniques of tree-ring dating. *Progress in Physical Geography: Earth and Environment*, 4(2), 161–188. <https://doi.org/10.1177/030913338000400202>
- Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., & Friesz, A. (2016). Modeled historical land use and land cover for the conterminous United States. *Journal of Land Use Science*, 11(4), 476–499. <https://doi.org/10.1080/1747423X.2016.1147619>
- Sovic, M., Fries, A., Martin, S. A., & Lisle Gibbs, H. (2019). Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, *Sistrurus catenatus*. *Evolutionary Applications*, 12(4), 664–678. <https://doi.org/10.1111/eva.12731>
- Stedman, A. L., Jaeger, C. P., Hileman, E. T., Jellen, B. C., Phillips, A., Swanson, B. J., & King, R. B. (2016). Multiple paternity in three wild populations of eastern massasauga (*Sistrurus catenatus*). *Herpetological Conservation and Biology*, 11(1), 160–167.

- Steen, D. A., Osborne, P. A., Dov, M., Patrick, D. A., & Gibbs, J. P. (2015). A preliminary investigation into the short-term effects of a prescribed fire on habitat quality for a snake assemblage. *Herpetological Conservation and Biology*.
- Szymanski, J., Pollack, C., Ragan, L., Redmer, M., Clemency, L., Voorhies, K., & Jaka, J. (2016). *Species Status Assessment for the Eastern Massasauga Rattlesnake (Sistrurus catenatus)*. US Fish and Wildlife Service.
- Tetzlaff, S. J., Ravesi, M. J., Allender, M. C., Carter, E. T., DeGregorio, B. A., Josimovich, J. M., & Kingsbury, B. A. (2017). Snake fungal disease affects behavior of free-ranging massasauga rattlesnakes (*Sistrurus catenatus*). *Herpetological Conservation and Biology*.
- Thünken, T., Bakker, T. C. M., Baldauf, S. A., & Kullmann, H. (2007). Active inbreeding in a cichlid fish and its adaptive significance. *Current Biology*, *17*(3), 225–229. <https://doi.org/10.1016/j.cub.2006.11.053>
- Toczydlowski, R. H., & Waller, D. M. (2023). Failure to purge: Population and individual inbreeding effects on fitness across generations of wild *Impatiens capensis*. *Evolution*, *77*(6), 1315–1329. <https://doi.org/10.1093/evolut/qpaa047>
- Trakhtenbrot, A., Nathan, R., Perry, G., & Richardson, D. M. (2005). The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions*, *11*(2), 173–181. <https://doi.org/10.1111/j.1366-9516.2005.00156.x>
- Turner, T. F., Osborne, M. J., Moyer, G. R., Benavides, M. A., & Alò, D. (2006). Life history and environmental variation interact to determine effective population to census size ratio. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1605), 3065–3073. <https://doi.org/10.1098/rspb.2006.3677>
- Vickers, T. W., Sanchez, J. N., Johnson, C. K., Morrison, S. A., Botta, R., Smith, T., Cohen, B. S., Huber, P. R., Ernest, H. B., & Boyce, W. M. (2015). Survival and mortality of Pumas (*Puma concolor*) in a fragmented, urbanizing landscape. *PLOS ONE*, *10*(7), e0131490. <https://doi.org/10.1371/journal.pone.0131490>
- Vieira, F. G., Lassalle, F., Korneliussen, T. S., & Fumagalli, M. (2016). Improving the estimation of genetic distances from Next-Generation Sequencing data. *Biological Journal of the Linnean Society*, *117*(1), 139–149. <https://doi.org/10.1111/bij.12511>
- von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws ii). *Human Biology*, *10*(2), 181–213.
- Vranckx, G., Jacquemyn, H., Muys, B., & Honnay, O. (2012). Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology*, *26*(2), 228–237. <https://doi.org/10.1111/j.1523-1739.2011.01778.x>
- Waples, R. K., Larson, W. A., & Waples, R. S. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity*, *117*(4), 233–240.

- Waples, R. S. (2016). Making sense of genetic estimates of effective population size. *Molecular Ecology*, 25(19), 4689–4691. <https://doi.org/10.1111/mec.13814>
- Waples, R. S. (2022). What is Ne, anyway? *Journal of Heredity*, 113(4), 371–379. <https://doi.org/10.1093/jhered/esac023>
- Waples, R. S. (2024). The Ne/N ratio in applied conservation. *Evolutionary Applications*, 17(5), e13695. <https://doi.org/10.1111/eva.13695>
- Waples, R. S., & Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, 8(4), 753–756. <https://doi.org/10.1111/j.1755-0998.2007.02061.x>
- Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7(2), 256–276. <https://doi.org/10.1039/b316709g>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370. <https://doi.org/10.2307/2408641>
- Wilder, S. M., Raubenheimer, D., & Simpson, S. J. (2016). Moving beyond body condition indices as an estimate of fitness in ecological and evolutionary studies. *Functional Ecology*, 30(1), 108–115. <https://doi.org/10.1111/1365-2435.12460>
- Willi, Y., Buskirk, J. V., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, 37(Volume 37, 2006), 433–458. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110145>
- Wilson, D. S., Tracy, C. R., & Tracy, C. R. (2003). Estimating age of turtles from growth rings: A critical evaluation of the technique. *Herpetologica*, 59(2), 178–194. [https://doi.org/10.1655/0018-0831\(2003\)059\[0178:EAOTFG\]2.0.CO;2](https://doi.org/10.1655/0018-0831(2003)059[0178:EAOTFG]2.0.CO;2)
- Wooldridge, B., Orland, C., Enbody, E., Escalona, M., Mirchandani, C., Corbett-Detig, R., Kapp, J. D., Fletcher, N., Ammann, K., Raimondi, P., & Shapiro, B. (2024). *Limited genomic signatures of population collapse in the critically endangered black abalone (Haliotis cracherodii)* (p. 2024.01.26.577275). bioRxiv. <https://doi.org/10.1101/2024.01.26.577275>
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16(2), 97–159.
- Wright, S. (1943). Isolation by Distance. *Genetics*, 28(2), 114–138.
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1), 76–82.
- Yineger, H., Schmidt, D. J., & Hughes, J. M. (2014). Genetic structuring of remnant forest patches in an endangered medicinal tree in North-western Ethiopia. *BMC Genetics*, 15(1), 31. <https://doi.org/10.1186/1471-2156-15-31>

## APPENDIX A:

### Supporting Information: Pitfalls and windfalls of detecting demographic declines using population genetics in long-lived species



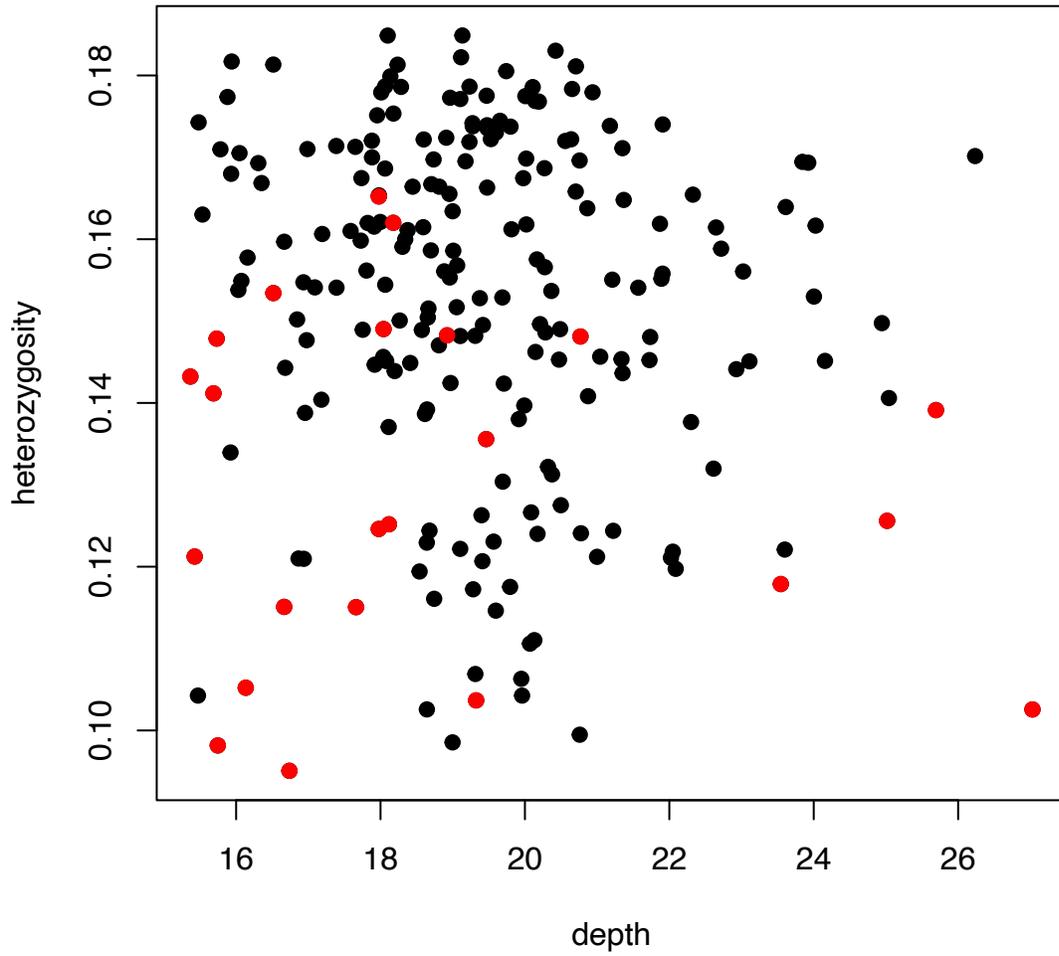
**Figure S2.1.** Plots show percent detection in  $\theta_W$  and  $\pi$  between oldest and youngest individuals for perennial models ( $A = 2, 5, 10, 20$ ). The old age bin is made up of individuals whose age at a given timepoint fell above the 50<sup>th</sup> quantile. The young age bin is made up of the youngest individuals at a given timepoint. Sample size is the same between bins. Percent detection is defined as the percent of replicates where we found a significant difference between the old and young age bins. The red dashed line indicates when the bottleneck event occurred. The gray shaded portions of the plot indicate when samples were taken every 50 ticks of the simulation, where the unshaded portion of the plot indicates when samples were taking every five ticks of the simulation.

**Table S2.1.** Mean sample sizes for annual and perennial simulations for age and temporal analyses. Values in parentheses represent the ranges in the data across replicates and timepoints. Annual sample sizes were calculated as the mean sample size for age sampling across perennial simulations.

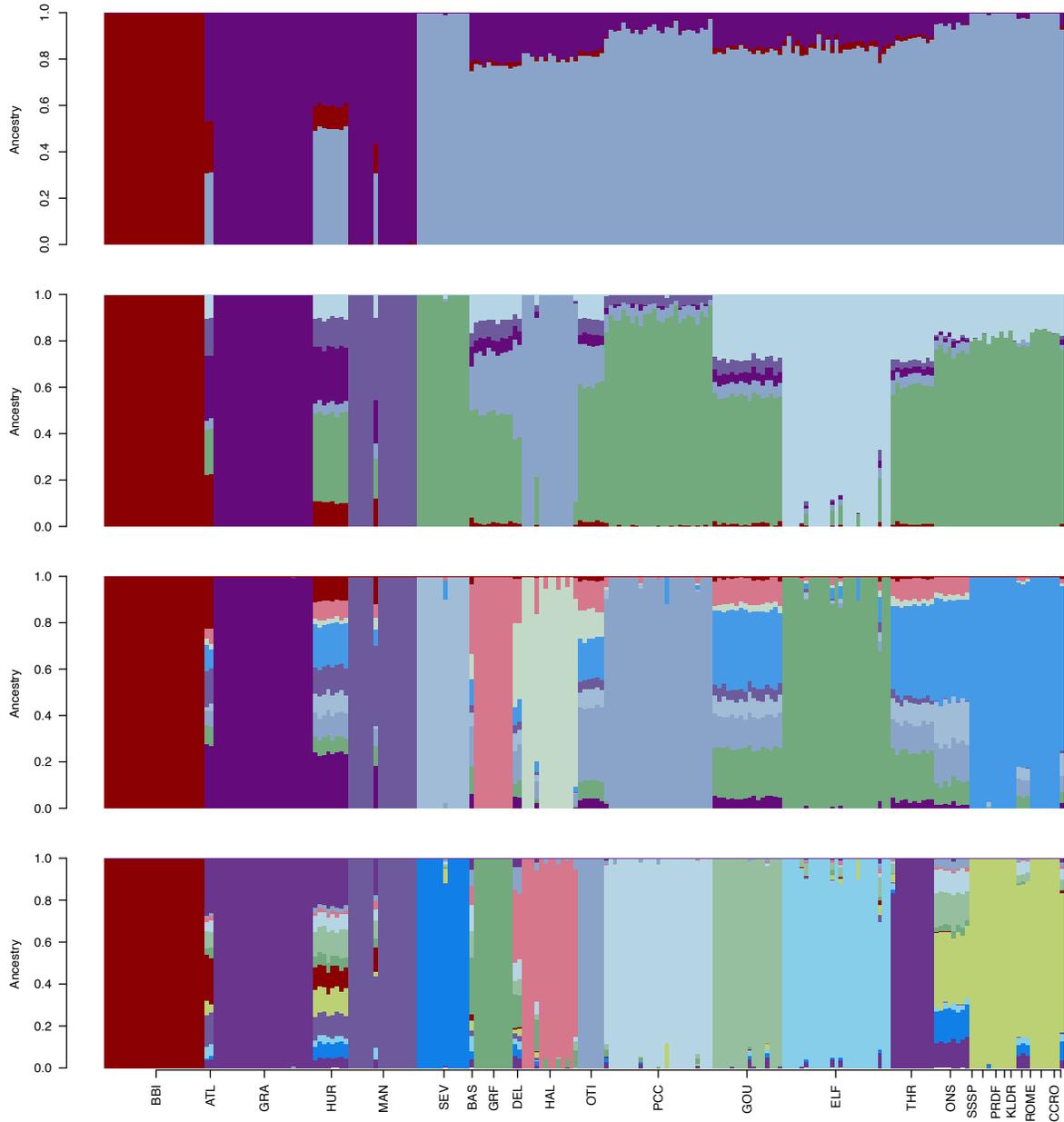
		Bottleneck intensities		
		2	10	100
Average age	1	113 (111, 116)	113 (111, 116)	106 (103, 116)
	2	132 (128, 134)	132 (129, 134)	115 (109, 136)
	5	112 (110, 115)	113 (110, 115)	106 (103, 114)
	10	104 (99, 112)	104 (100, 114)	103 (96, 113)
	20	104 (98, 108)	104 (99, 109)	102 (98, 106)

APPENDIX B:

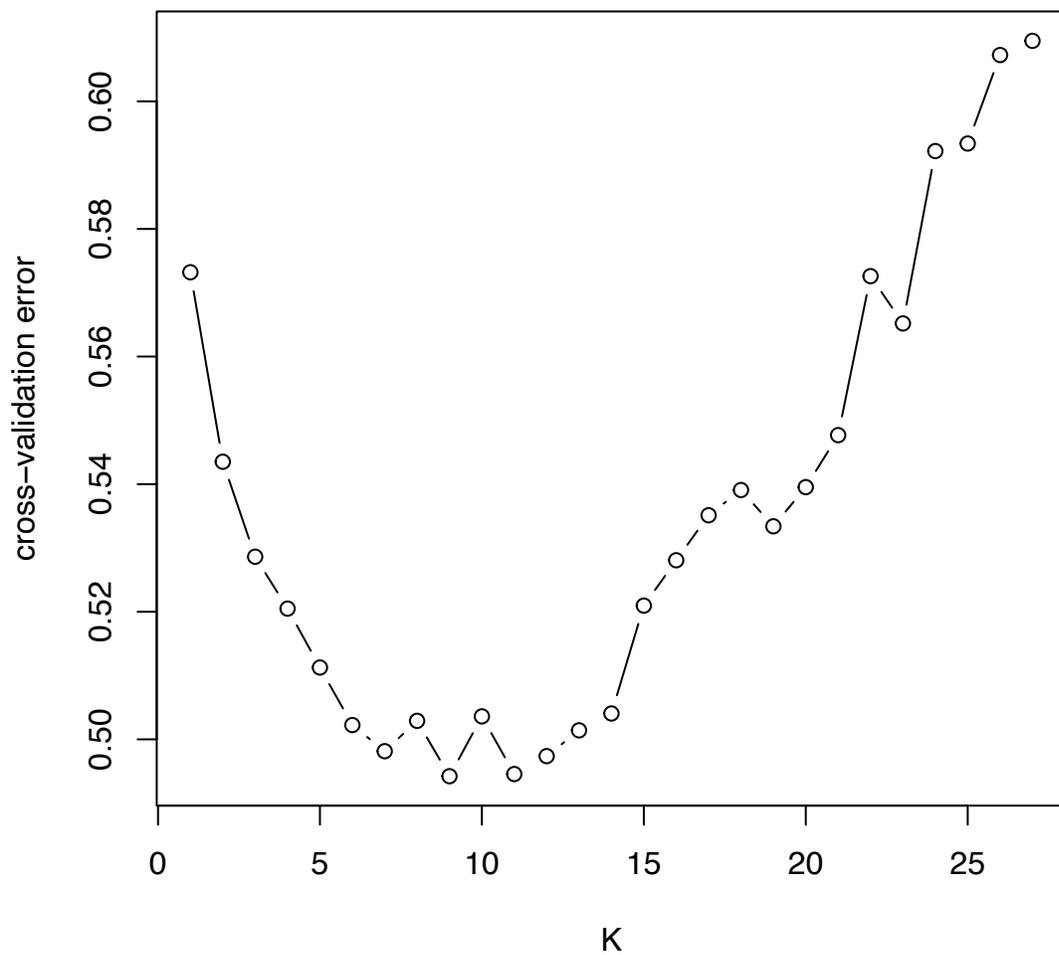
Supporting Information: Anthropogenic land use changes associated with increased inbreeding in  
a threatened rattlesnake



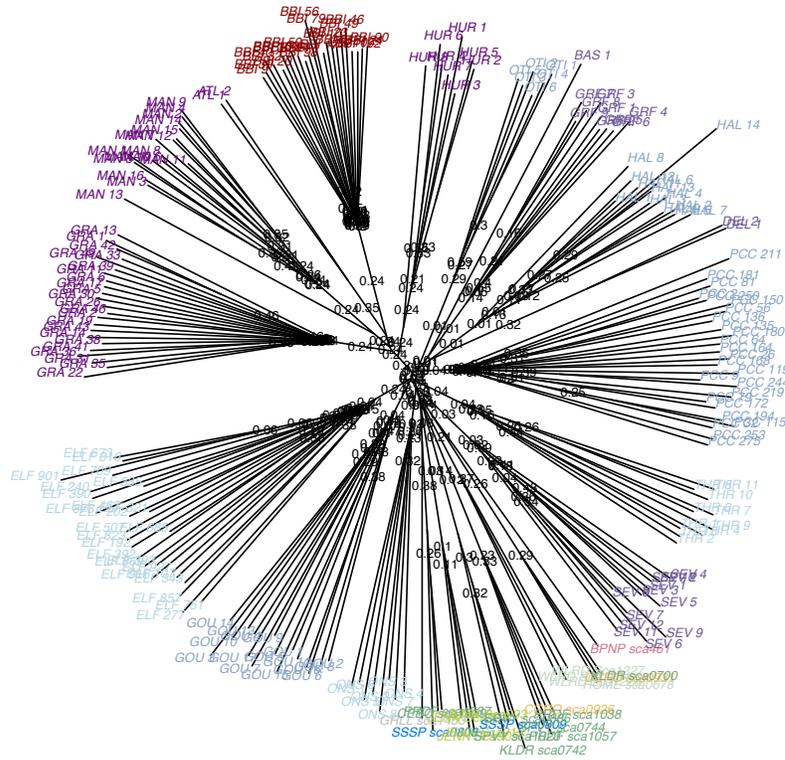
**Figure S3.1.** Plot shows individual depth vs heterozygosity for eastern massasauga whole genome sequences. Red points represent sequences from Mathur et al. 2023, where as black points are from this study.



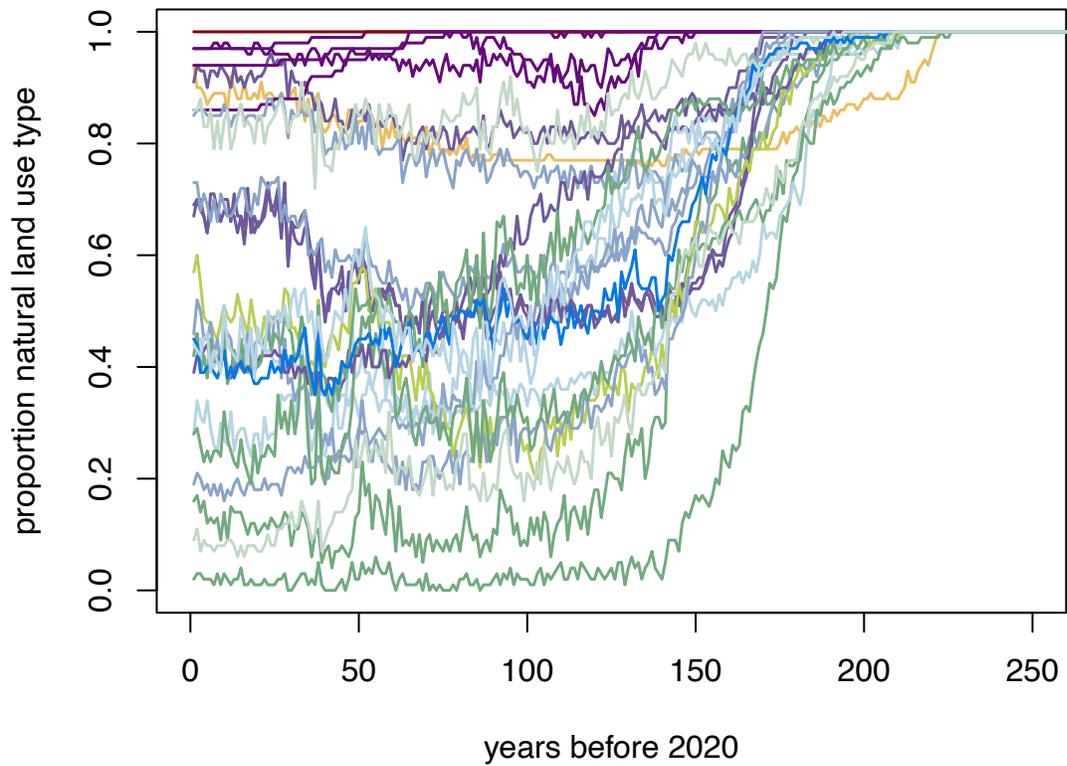
**Figure S3.2.** Genetic clustering from ADMIXTURE. Each bar represents an individual, and the color of the bar corresponds to its assignment to a genetic cluster.



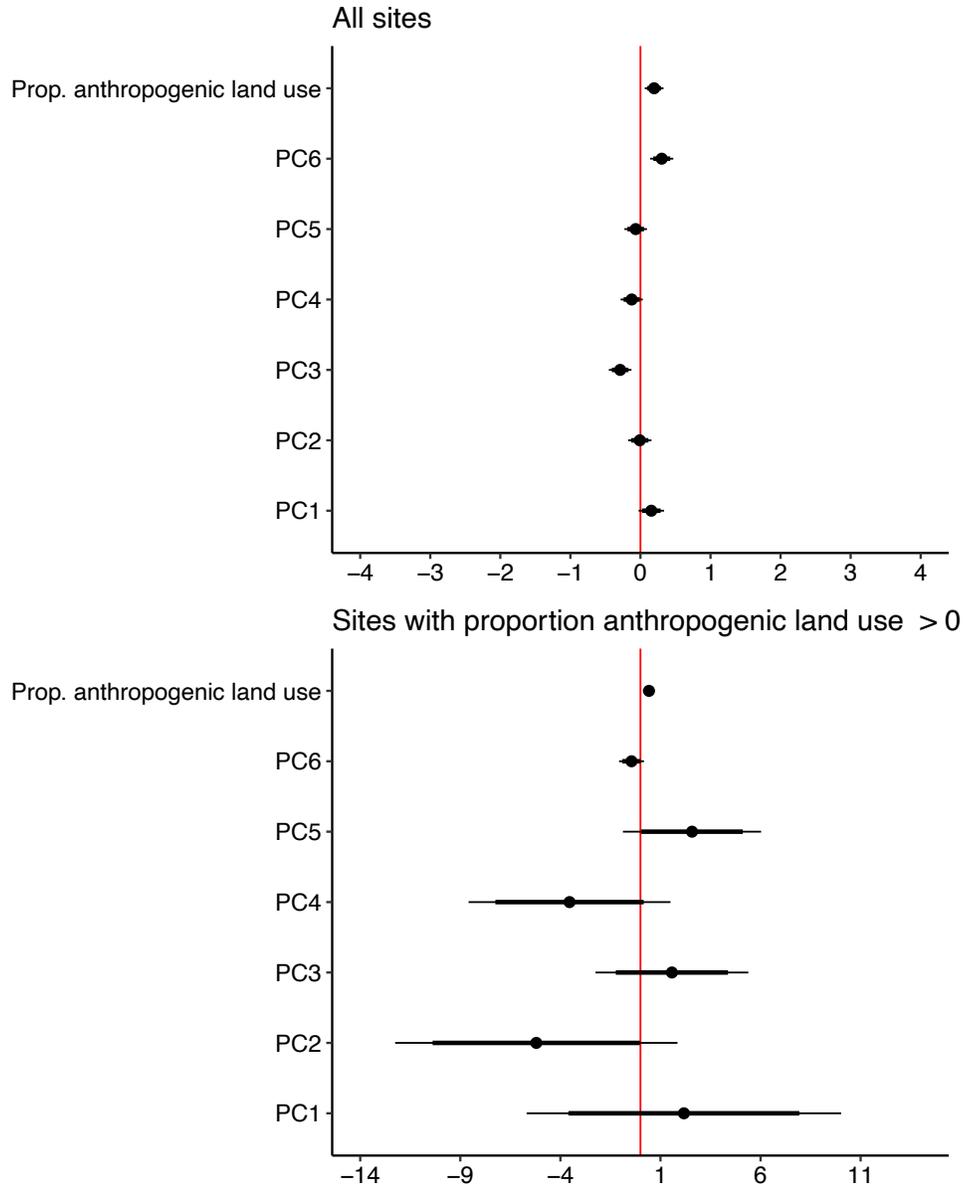
**Figure S3.3.** Cross-validation error from ADMIXTURE.  $K = 9$  has the lowest cross-validation error.



**Figure S3.4.** Neighbor-joining tree built using genetic distances between eastern massasaugas from 26 sites. Color of sample name corresponds to geographic regions in Figure 4.1.



**Figure S3.5.** Proportion of natural land use types within a 10x10 km grid centered on a focal eastern massasauga rattlesnake population over the past 250 years. Land use/land cover data from Li. et al. 2023.



**Figure S3.6.** Effect sizes (circles) and their associated 85% (thick bars) and 95% (thin bars) confidence intervals based on the top-ranked linear models for all populations (top) and populations with average proportion of anthropogenic land use greater than zero.

## APPENDIX C:

Supporting Information: Inbreeding reduces fitness in spatially structured populations of a threatened rattlesnake

### **Supplemental Methods**

#### *Capture-recapture surveys*

We conducted field surveys during the active season from 2009 through 2023. Data was collected from Cass County for 14 years (2009-2023, no 2020) and from Barry County for 13 years (2011-2023). After recording location information, captured snakes were brought back to the lab where we recorded sex (Schaefer, 1934), snout-vent length (SVL (cm), (Quinn & Jones, 1974), and other relevant morphometrics. Snakes were individually marked with passive integrated transponder (PIT) tags. We collected blood from the caudal vein for genetic analyses and released individuals at their site of capture. 30 gravid females at the Cass County site were held in the lab until partition as part of a separate study (Hileman et al., 2015; Stedman et al., 2016). We included these individuals and their offspring to verify our pedigree reconstruction.

#### *Initial library preparation*

To identify polymorphic loci in our focal populations, 64 individuals were included in an initial RAD library (39 from Barry County; 25 from Cass County) (Ali et al 2016). We submitted a single library for sequencing with paired-end 150 bp reads on an Illumina HiSeq 4000 at Michigan State University's genomic core facility. Sequencing yielded 310 million sequence reads and quality checks revealed high Q-scores (85% > Q30), indicating this library was successful.

#### *Initial library bioinformatics*

Raw sequence reads from the initial RADseq library were processed through two separate

bioinformatic pipelines: Stacks (Catchen et al., 2013) and ipyrad (Eaton 2014). Reads were filtered for quality and clustered both within and among individuals in order to identify sequences containing single nucleotide polymorphisms (SNPs). For the ipyrad analysis, reads were mapped to the draft genome for the eastern massasauga rattlesnake (Mathur et al. 2023). Mapping reads to a reference genome improves genotyping accuracy and facilitates a wider set of analyses that take advantage of precise genomic location information. Average read depth across all samples for sequences mapped to the reference genome was 15X and in total 11,862 individual loci containing SNPs were identified.

To identify target SNPs for RAD-based capture assays, 64 individuals were included in an initial BestRAD library (39 from Barry County and 25 from Cass County) . We sequenced the library on an Illumina HiSeq 4000 with 150 bp paired-end reads at the Michigan State University Core Genomics Facility. This initial library was used to identify 3,391 putative neutral candidate SNPs with minor allele frequencies  $> 0.05$  in both populations, spaced at least 100 kbp apart on the genome. Custom baits to target these SNPs were designed by Arbor Biosciences (Ann Arbor, MI).

#### *Bait design*

An initial candidate set of loci was generated from the 64-individual pilot data set that included SNPs proximate to venom-related genes (*svmp*, *svsp*, *pla2*, and *myotoxin*) and immune-related genes (*MHCI*, *MHCII*, and *defensin*) as well as loci with no association with known functional genes. From the SNPs located within loci with no known functional role, a subset of candidate SNPs (3,391 putatively neutrally-evolving SNPs) with known positions on the draft reference genome were used to design baits for sequence capture, retaining baits with fewer highly repetitive sequences and more unique BLAST hits based on criteria previously used by

Arbor Biosciences. Baits were designed in sets of three eighty-base oligonucleotides tiled along an approximately 120-base RAD locus to maximize capture probability and efficiency.

### *RAPTURE genotyping*

We extracted DNA from all unique blood samples up until 2019 at Cass County and 2021 at Barry County using Qiagen DNeasy Blood and Tissue kits. DNA was quantified using a Qubit 2.0 and quality was assessed using agarose gels. Twelve BestRAD libraries were prepared containing 1,056 individuals (30 individuals sequenced as part of multiple libraries to act as technical duplicates) (Ali et al., 2016). Libraries were pooled for bait capture, resulting in three final pooled libraries. To validate even representation of individual libraries within pools, we first sequenced the pooled libraries on one lane of a MiSeq v2 nano. We ran libraries on a single lane of a NovaSeq 6000 S4 flow cell at the Michigan State University Core Genomics Facility.

We demultiplexed sequence data using STACKS v. 2.59 (Catchen et al., 2013) and aligned them to the eastern massasauga reference genome (Mathur et al., 2023) using BWA v. 07.17 (Li, 2013). We used Samtools v. 1.9 (Li et al., 2009) to index and sort alignments, as well as filter out unpaired reads. We used the Stacks ref pipeline to call SNPs and remove PCR duplicates using the following parameters: `--rm-pcr-duplicates -X "populations: -p 2 -r 0.75"`. We filtered SNPs using bcftools v.1.9.64 (Li, 2011) to retain SNPs with more than seven reads in 90% of individuals and with genotype quality scores greater than 19 in 90% of individuals. We filtered out SNPs with excess heterozygosity ( $p < 0.05$ ) using the "HWExactStats" function in the HardyWeinberg R package v. 1.7.5 (Graffelman, 2015). We also filtered to retain one randomly selected bi-allelic SNP per targeted genomic region. Finally, using custom R code, we removed SNPs with more than 10% missing data, and removed individuals with greater than 20% missing genotype calls or that had uneven distributions of reference and alternate allele read

counts indicative of sample contamination (R Core Team 2023). For principal components analysis, we filtered out SNPs with a minor allele frequency below 0.05. For pedigree reconstruction, we filtered out SNPs with a minor allele frequency less than 0.1 separately for each site. We calculated effective population size ( $N_e$ ) for each population using the strataG package in R, using linkage disequilibrium based on Pearson correlation approximation as in Waples et al. (2016) (Archer et al., 2017; Waples et al., 2016). We estimated 95% confidence intervals around  $N_e$  estimates using jackknife resampling.

### *Inference of individual birth years*

To estimate individual birth years, we predicted individual SVL with the von Bertalanffy (Fabens) top model (Table S4.3). Specifically, we predicted SVL backward in time for individuals first captured with an SVL < asymptotic size, assuming a minimum size of 12 cm, the size of the smallest neonate observed at either site. We refined the backward predictions of SVL (also needed to estimate longevity; see below) based on our best estimates and knowledge of when individuals were age zero. To do this, we used the following rules: (1) If an individual was captured at age zero, we used the recorded (measured) SVL at that capture event. (2) If an individual was not captured at age zero, we evaluated their predicted SVL in the year prior to the year that their first predicted or recorded SVL was > 24.5 cm (Cass) or > 21.8 cm (Barry). These sizes represent the maximum SVLs of known age zero captures; thus, we assumed that an SVL > 24.5 cm (Cass) or > 21.8 cm (Barry) indicated an individual was age 1 or older. If the predicted SVL in the year prior was less than the mean SVL at age zero (19.2 cm for Cass and 19.3 cm for Barry), then we increased the individual's age zero SVL to the mean. If an individual was not captured at age zero, but their predicted SVL at age zero was > the mean size at age zero, then we kept the predicted age zero size. We also checked for individuals that were known to be age

one at their first capture and were smaller than the maximum SVL of known age zero captures. We assigned these individuals the mean size at age zero at the occasion prior to their first capture. This correction ensured that individuals known to be age one were not incorrectly treated as age zero at age one. We used the predicted year an individual was age zero as their estimated birth year.

#### *Calculation of longevity and years contributing offspring to the pedigree*

To account for differences in survival and age at the end of the study when assessing the impact of  $F_{\text{grm}}$  on reproductive output, we calculated the number of years an individual could have contributed offspring to the pedigree. We used the top ranked CJS model to predict yearly cumulative survival probabilities based on SVL, site, and  $F_{\text{grm}}$  for every year starting in an individual's estimated birth year. We assumed an individual had died or left the population when its cumulative survival probability was less than 0.005. For individuals who were considered deceased before the last year new captures were included in pedigree reconstruction (2019 for ELF, and 2021 for PCCI), we calculated individual longevity (inferred death year - inferred birth year). For individuals who were considered alive during the last year individuals were included in pedigree reconstruction, we used that year as an upper cut off for when that individual could have contributed to the pedigree ("last year contributing", [2019 or 2021] – inferred death year). We do not account for any variation in number of years it took an individual to become sexually mature.

**Table S4.1.** Confidence probabilities for different categories of parent-offspring trios in the Barry County pedigree estimated by reconstructing pedigrees from genotype data simulated from the empirical reconstructed pedigree (N = 10 simulated pedigrees). Individuals of type “G” are genotyped, “D” are dummy individuals inferred to exist through pedigree relationships, and “X” indicates no individual assigned. Bolded values represent types of pedigree relationships that were used in analyses. Italicized values represent types of relationships that were dropped from analyses due to low confidence.

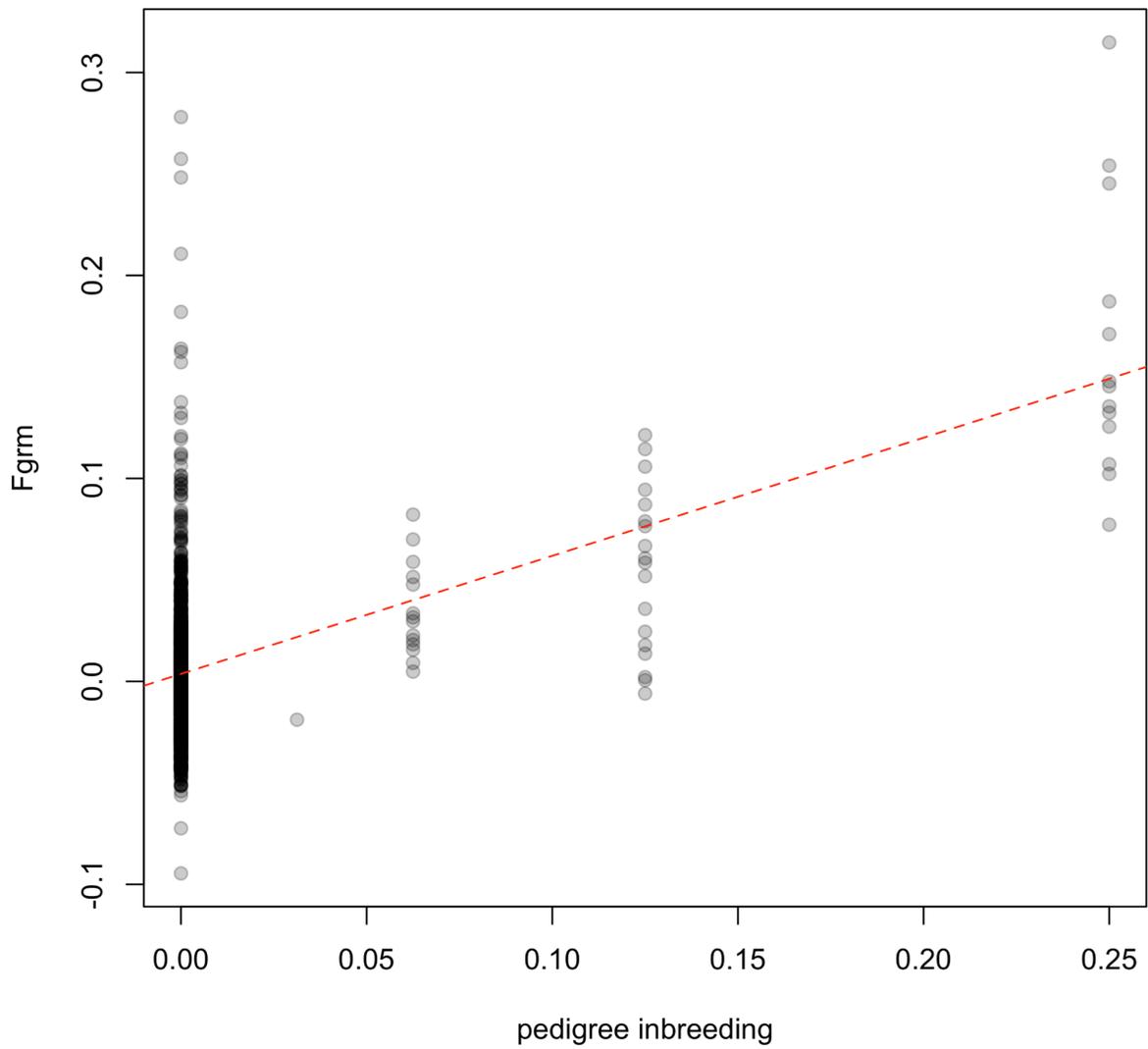
Focal type	Dam type	Sire time	Dam confidence probability	Sire. confidence probability	Pair confidence probability	N
G	<b>G</b>	<b>G</b>	<b>1.000</b>	<b>1.000</b>	1.000	765
G	<b>G</b>	D	<b>1.000</b>	1.000	1.000	450
G	<b>G</b>	X	<b>0.994</b>	NA	NA	166
G	D	<b>G</b>	NA	NA	NA	0
G	D	D	0.000	0.427	0.000	255
G	D	X	NA	NA	NA	0
G	X	<b>G</b>	NA	<b>1.000</b>	NA	42
G	X	D	NA	1.000	NA	9
G	X	X	NA	NA	NA	1173
D	<b>G</b>	<b>G</b>	<b>1.000</b>	<b>1.000</b>	1.000	15
D	<b>G</b>	D	<b>1.000</b>	1.000	1.000	7
D	<b>G</b>	X	<i>0.250</i>	NA	NA	4
D	D	<b>G</b>	NA	NA	NA	0
D	D	D	NA	NA	NA	0
D	D	X	NA	NA	NA	0
D	X	<b>G</b>	NA	NA	NA	0
D	X	D	NA	NA	NA	0
D	X	X	NA	NA	NA	325

**Table S4.2.** Confidence probabilities for different categories of parent-offspring trios in the Cass County pedigree estimated by reconstructing pedigrees from genotype data simulated from the empirical reconstructed pedigree ( $N = 10$  simulated pedigrees). Individuals of type “G” are genotyped, “D” are dummy individuals inferred to exist through pedigree relationships, and “X” indicates no individual assigned. Bolded values represent types of pedigree relationships that were used in analyses. Italicized values represent types of relationships that were dropped from analyses due to low confidence.

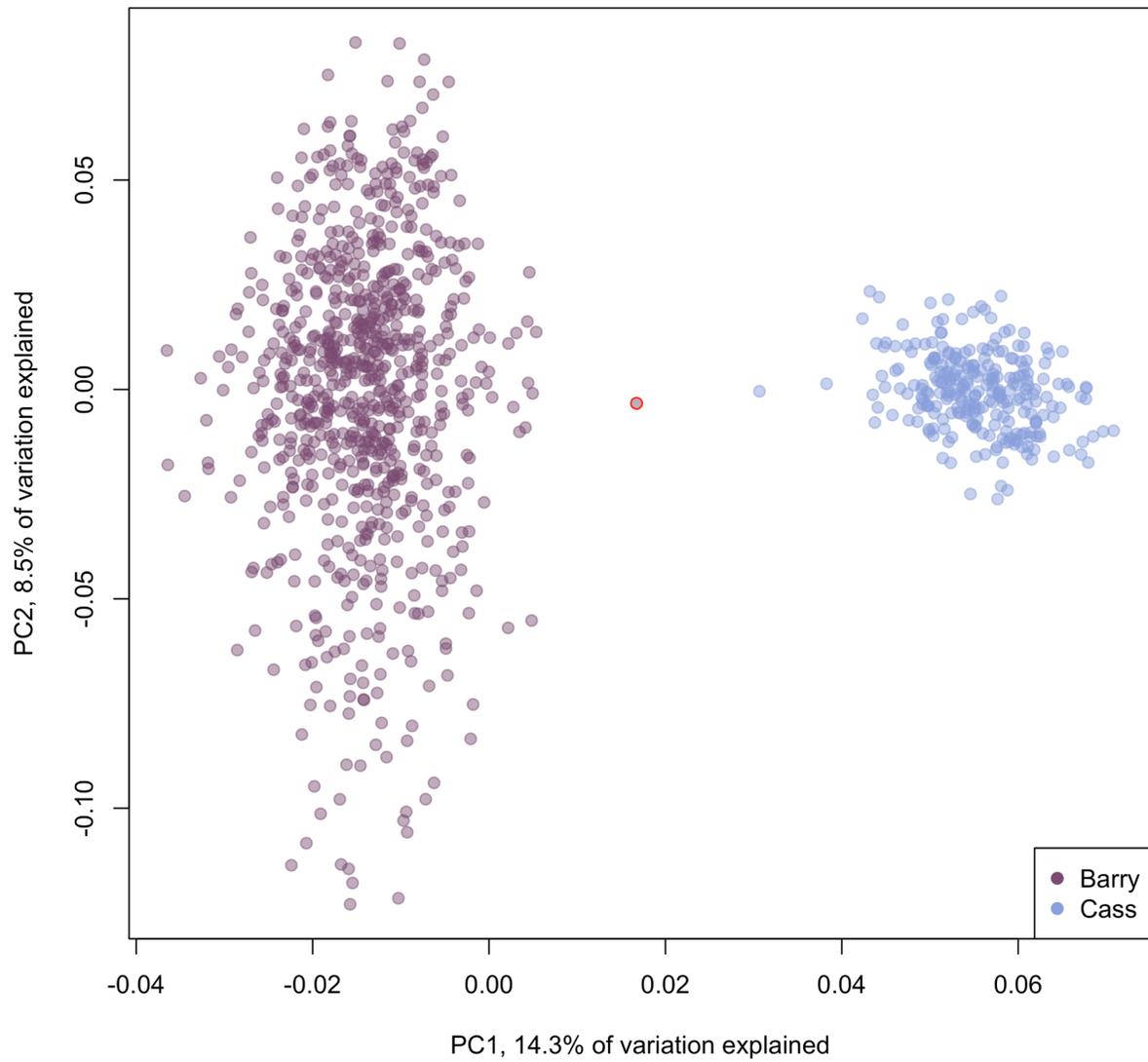
Focal type	Dam type	Sire time	Dam confidence probability	Sire. confidence probability	Pair confidence probability	N
G	<b>G</b>	<b>G</b>	<b>1.000</b>	<b>1.000</b>	1.000	3969
G	<b>G</b>	D	<b>1.000</b>	1.000	1.000	2564
G	<b>G</b>	X	<b>1.000</b>	NA	NA	166
G	D	<b>G</b>	0.000	<b>1.000</b>	0.000	2
G	D	D	0.000	0.506	0.000	308
G	D	X	NA	NA	NA	0
G	X	<b>G</b>	NA	<b>0.989</b>	NA	94
G	X	D	NA	0.911	NA	45
G	X	X	NA	NA	NA	1212
D	<b>G</b>	<b>G</b>	<b>0.986</b>	<b>0.986</b>	0.986	144
D	<b>G</b>	D	<b>1.000</b>	1.000	1.000	45
D	<b>G</b>	X	<i>0.842</i>	NA	NA	19
D	D	<b>G</b>	0.000	<i>0.000</i>	0.000	1
D	D	D	0.000	0.000	0.000	2
D	D	X	NA	NA	NA	0
D	X	<b>G</b>	NA	<b>0.950</b>	NA	20
D	X	D	NA	NA	NA	0
D	X	X	NA	NA	NA	392

**Table S4.3.** Candidate set of von Bertalanffy models (Fabens 1965) constructed from capture-recapture data collected from eastern massasauga populations (*Sistrurus catenatus*) in Cass and Barry counties between 2009–2023. Model parameters include asymptotic size ( $L_\infty$ ) and the growth constant ( $K$ ).  $\Delta AIC_c$  is the difference between the  $AIC_c$  of a given model and the top-ranked model;  $\omega_i$  is the probability that the model is the most parsimonious given the data and relative to other models in the candidate set;  $k$  is the number of parameters in the model;  $-2(\log)\mathcal{L}$  is the relative fit of the model.

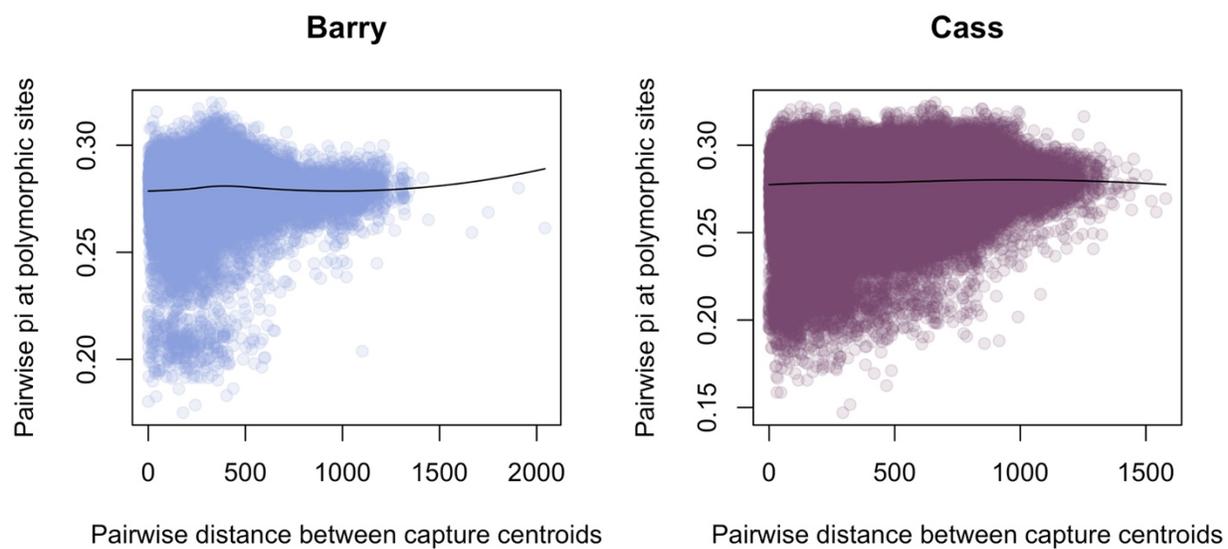
Models	$\Delta AIC_c$	$\omega_i$	$k$	$-2(\log)\mathcal{L}$
$L_\infty(\text{site}+\text{sex}), K(.)$	0	0.283	5	-1126.1
$L_\infty(\text{site}+\text{sex}), K(\text{sex})$	0.97	0.174	6	-1125.5
$L_\infty(\text{site}*\text{sex}), K(.)$	1.04	0.169	6	-1125.6
$L_\infty(\text{site}*\text{sex}), K(\text{sex})$	2.31	0.089	7	-1125.2
$L_\infty(\text{sex}), K(\text{site}+\text{sex})$	2.66	0.075	6	-1126.4
$L_\infty(\text{sex}), K(.)$	2.93	0.065	4	-1128.6
$L_\infty(\text{site}+\text{sex}), K(\text{site}+\text{sex})$	2.94	0.065	7	-1125.5
$L_\infty(\text{sex}), K(\text{sex})$	4.04	0.038	5	-1128.1
$L_\infty(\text{site}*\text{sex}), K(\text{site}+\text{sex})$	4.39	0.032	8	-1125.2
$L_\infty(\text{site}*\text{sex}), K(\text{site}*\text{sex})$	6.48	0.011	9	-1125.2



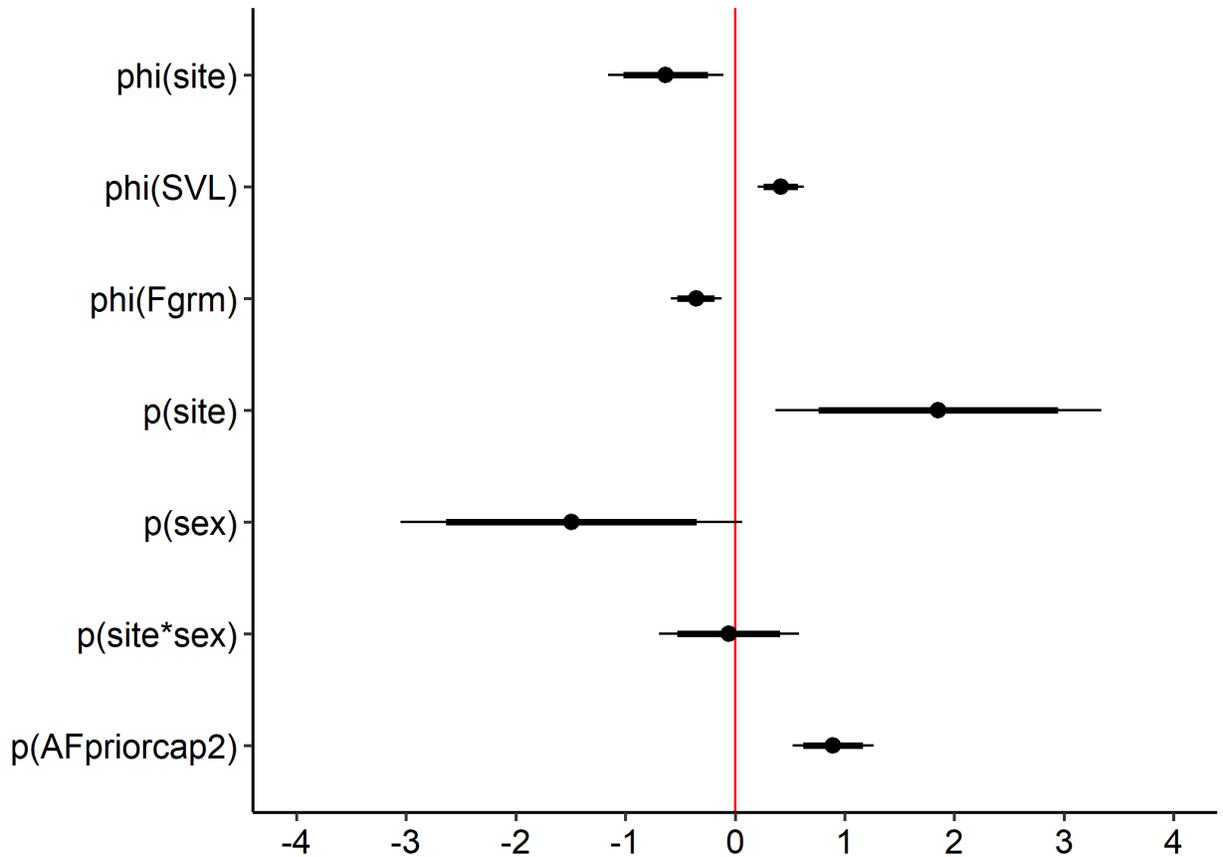
**Figure S4.1.** Plot shows the relationship between pedigree inbreeding and  $F_{grm}$  for eastern massasaugas at two sites in Michigan ( $r^2 = 0.216$ ,  $p < 2e^{-16}$ ).



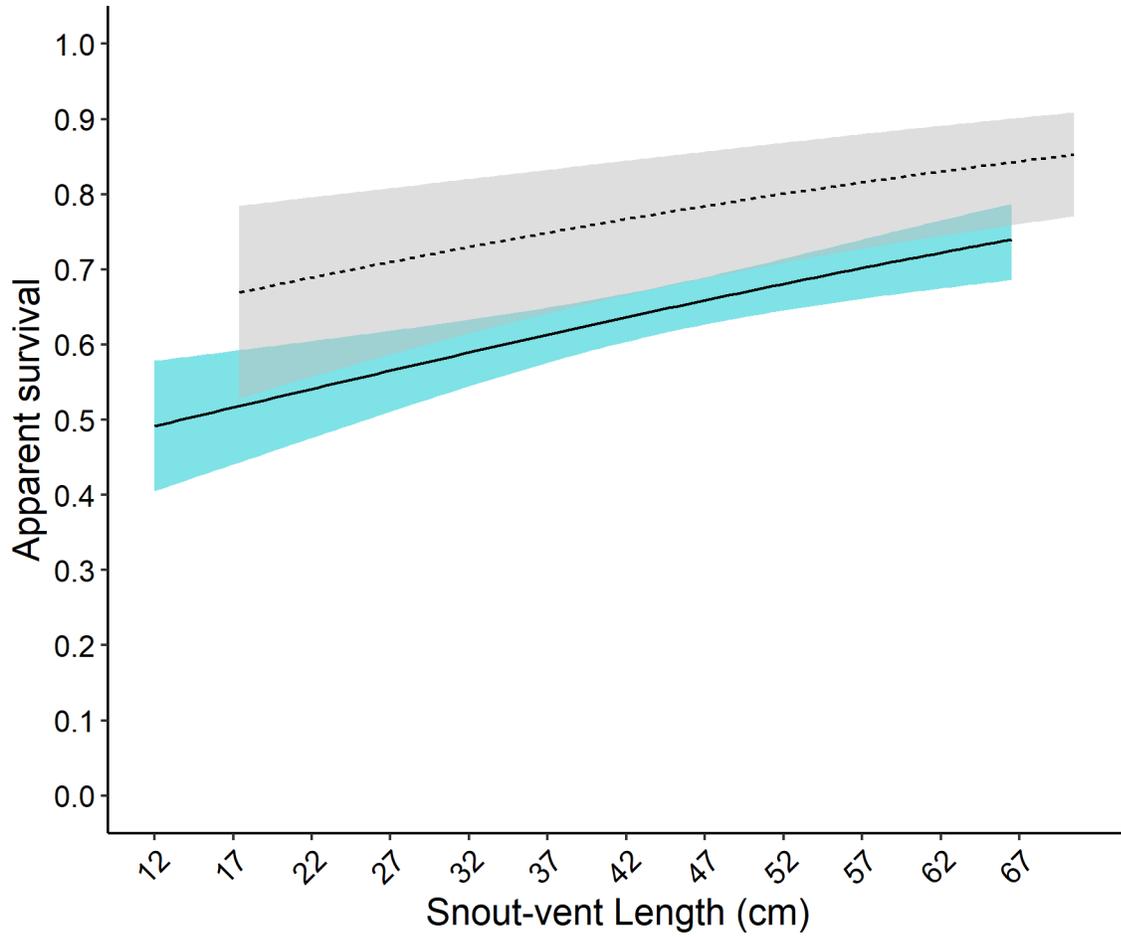
**Figure S4.2.** Principal component analysis of eastern massasaugas at two sites in Michigan. Individual circled in red is the potential migrant in Cass County that was removed from downstream analyses.



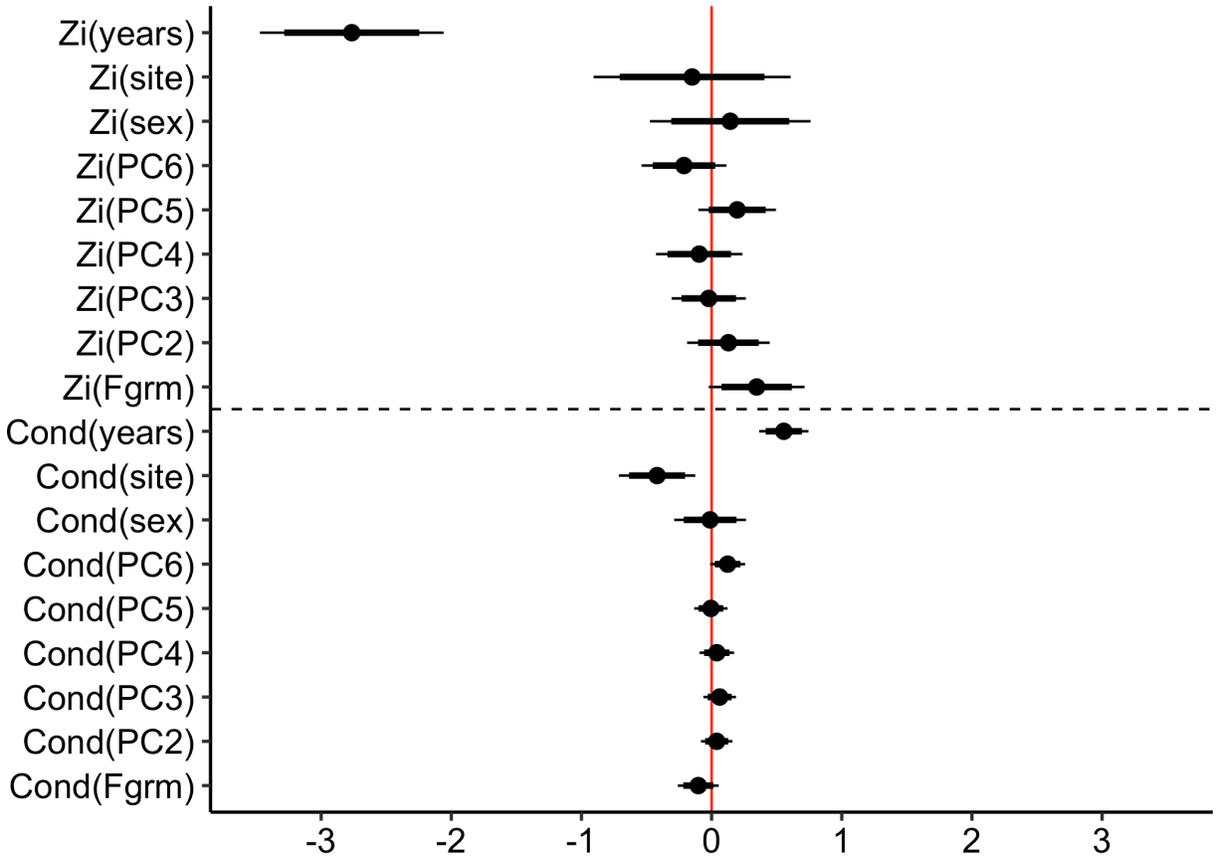
**Figure S4.3.** Plots show the relationship between pairwise  $\pi$  at polymorphic sites and pairwise distance between centroid locations for eastern massasaugas at two sites in Michigan, USA. Line represents smoothed loess line.



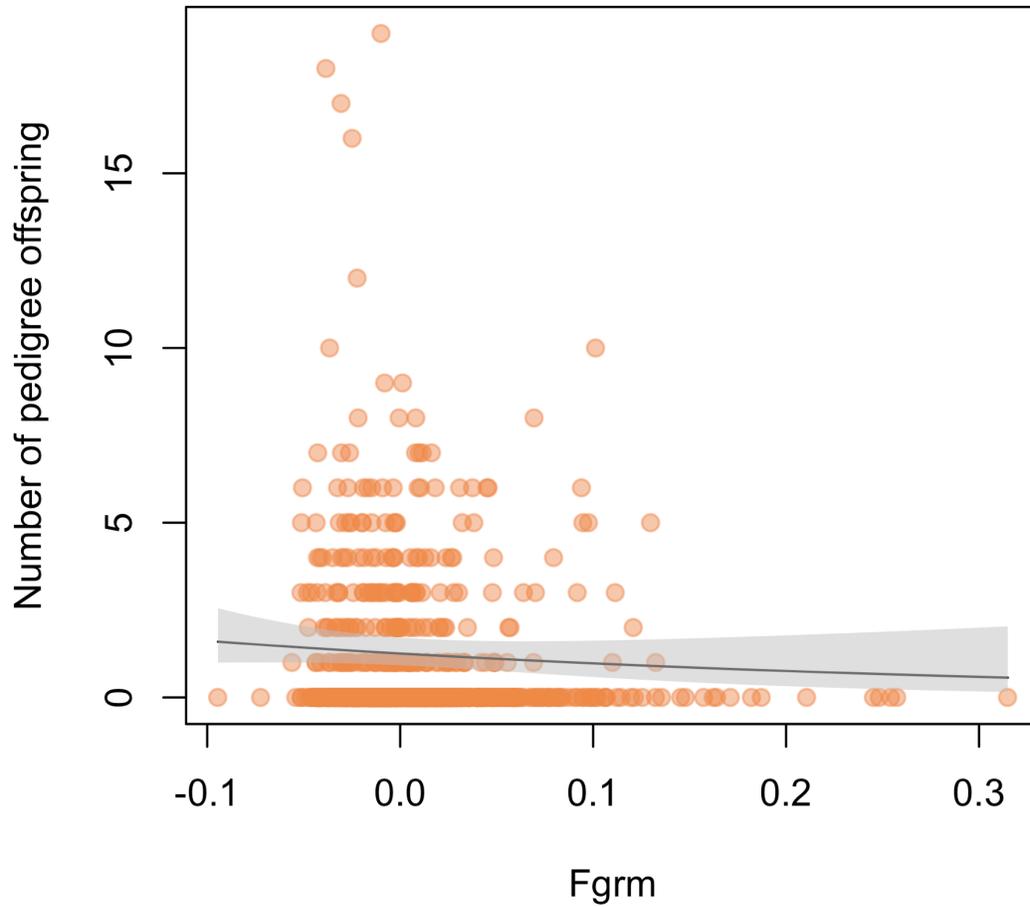
**Figure S4.4.** Standardized and zero-centered (red line) mean effect sizes (circles) of beta coefficients and their associated 85% (thick bars) and 95% (thin bars) confidence intervals based on the top-ranked Cormack-Jolly-Seber model. The model was fitted to capture-recapture data collected in Cass and Barry counties, MI, from 2009–2023. Shown are the effects of site (Cass), SVL, and  $F_{\text{grm}}$  on annual apparent survival and the effects of site (Cass), sex (male), and a (biennial) behavioral response to recapture for adult females (AFpriorcap2). Site (Barry) and sex (female), not shown, are the intercepts for survival and recapture probabilities, respectively.



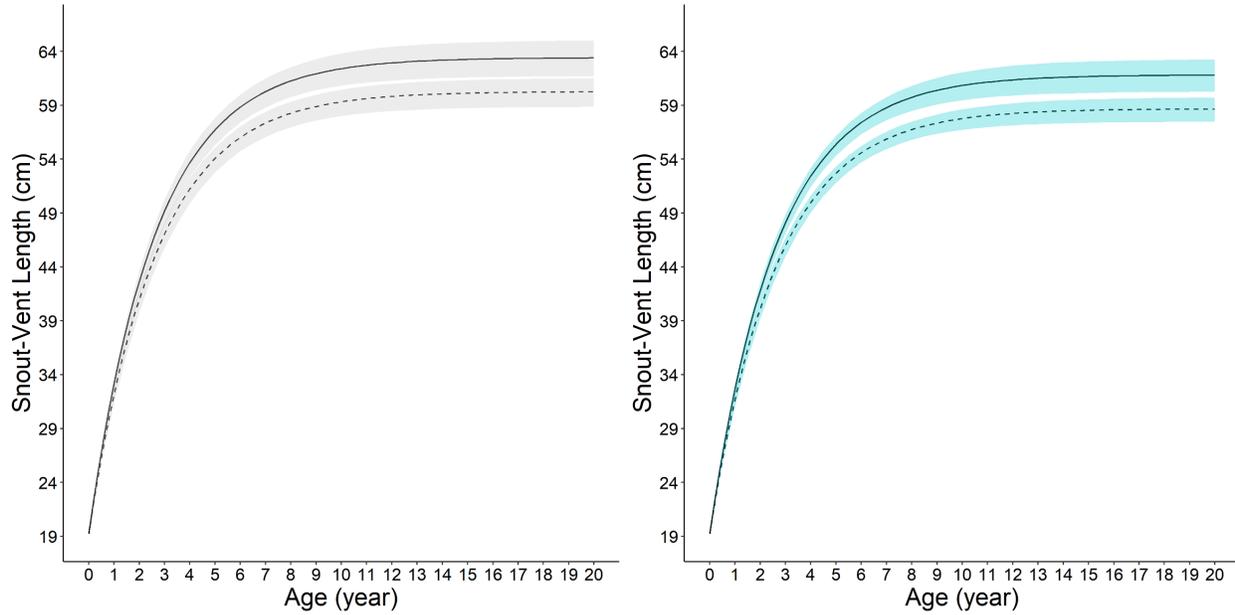
**Figure S4.5.** Effects of snout-vent length on annual apparent survival for eastern massasaugas (*Sistrurus catenatus*) from Barry (dashed line) and Cass (solid line) counties. Shaded bands represent 95% confidence intervals. To facilitate site comparisons, variation in  $F_{\text{grm}}$  was controlled for by holding it at its mean 0.00728.



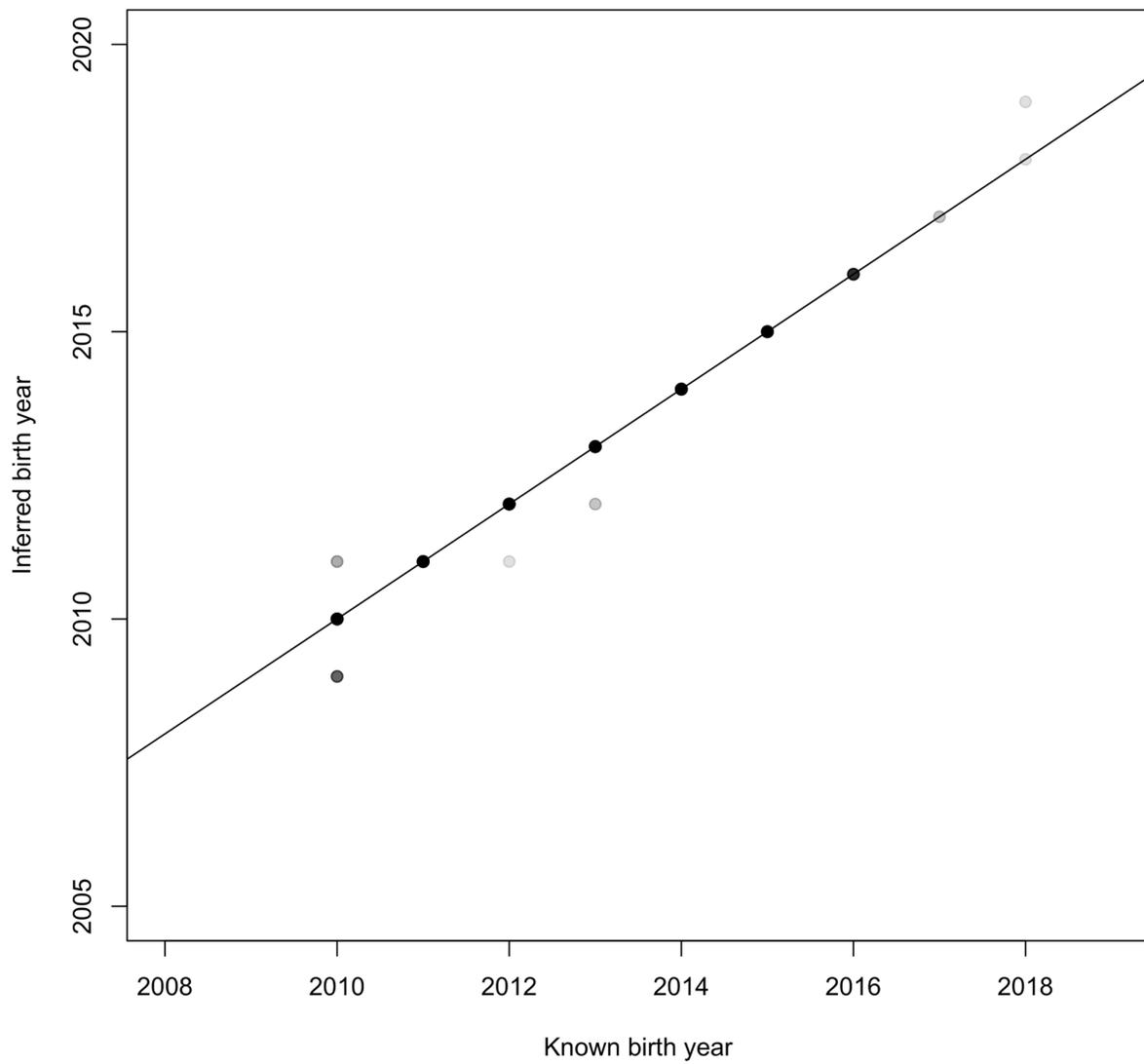
**Figure S4.6.** Standardized effect sizes (circles) and their associated 85% (thick bars) and 95% (thin bars) confidence intervals based on the top-ranked zero-inflated negative binomial model. Effect sizes above the dashed line are for the zero-inflated component of the model, and effect sizes below the dashed line are for the conditional component of the model. Shown are the effects of site (Barry), sex (males), years an individual contributed to the pedigree, genetic principal components 2-6, and  $F_{grm}$  on number of pedigree offspring.



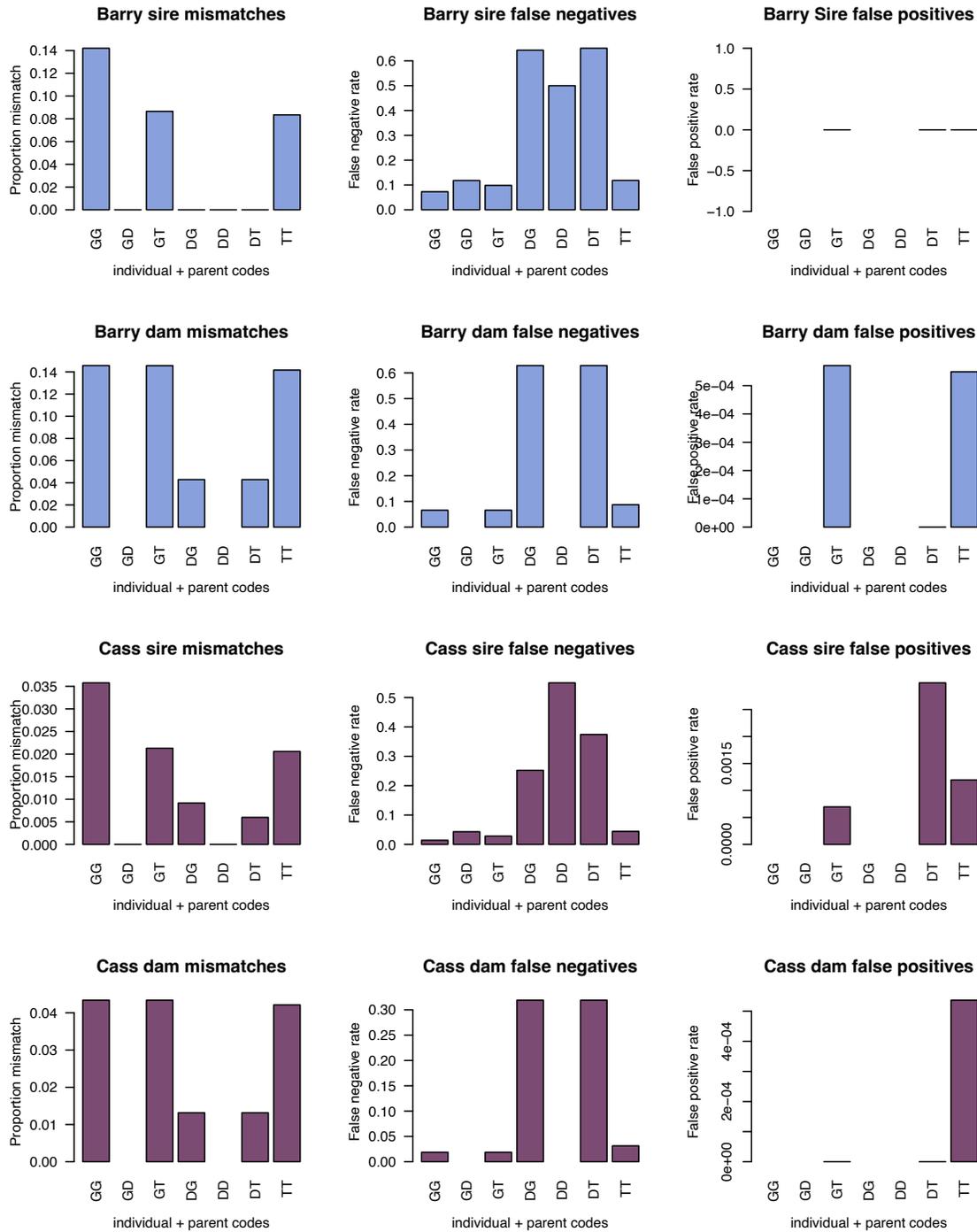
**Figure S4.7.** Effect of  $F_{grm}$  on number of offspring assigned in the pedigree. The solid line represents the predicted number of offspring. The shaded band represents the 95% confidence interval. Orange points are the data used to fit the model.



**Figure S4.8.** Eastern massasauga (*Sistrurus catenatus*) growth and asymptotic size for males (solid line) and females (dashed). Shaded bands represent 95% confidence intervals for Barry (left) and Cass (right) counties. Growth was constant between sexes and populations ( $K = 0.378$ , 95% CI = 0.349– 0.408), whereas asymptotic size was larger for males and females in Barry County. Size at birth (age 0 = 19.24 cm) was calculated from the average neonate snout-vent-length in Cass County ( $n = 412$ ).



**Figure S4.9.** Plot shows the correlation between known individual birth year and birth years inferred using snout-vent-length (SVL) and von Bertalanffy models.



**Figure S4.10.** Pedigree errors between the empirical pedigree and pedigrees reconstructed using genotype data simulated from the empirical pedigree ( $N = 10$  simulated pedigrees). Mismatches represent cases when a different parent was assigned to an individual in a simulated versus empirical pedigree. The false negative rate is the rate at which a parent was assigned in the empirical pedigree but was not in the simulated pedigree. The false positive (rate is the rate at which a parent was assigned in the simulated pedigree, but not assigned in the empirical pedigree).