Genetic Diversity, Inbreeding and Diet Variation in an Endangered Rattlesnake, the Eastern Massasauga (*Sistrurus c. catenatus*)

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

James Ernest Chiucchi Jr, M.S.

Graduate Program in Evolution, Ecology, and Organismal Biology

The Ohio State University

2011

*****

Dissertation Committee:

Dr. H. Lisle Gibbs, Advisor

Dr. Peter Curtis

Dr. Paul A. Fuerst
Copyright by

James Ernest Chiucchi Jr.

2011
Abstract

Many evolutionary and ecological factors can have strong and significant effects in small isolated populations when compared to larger, more connected populations. For example, isolated populations of endangered species are often at greater risk for extinction because of the potential loss of genetic diversity and the negative effects of inbreeding and inbreeding depression. Reduced levels of gene flow among isolated populations may also lead to differentiation in important ecological factors such as diet. The eastern massasauga (*Sistrurus c. catenatus*) is an endangered snake that is found in isolated populations throughout eastern North America. In this dissertation, I address both genetic and ecological consequences of these snakes existing in small isolated populations.

In Chapter 2, I used data from 19 microsatellite loci to assess levels of genetic variation and to compare rates of historical and contemporary gene flow among populations of the these snakes. Overall, massasauga populations display high levels of genetic differentiation and differ by an order of magnitude in genetic effective population size. Historical and contemporary migration rates are low and similar in magnitude. There is little evidence that populations have experienced a sharp decline in population size and a comparative modeling approach favors a model of long-term drift-migration equilibrium.
In Chapter 3, levels of inbreeding were accessed in these populations and I tested for the presence of a heterozygosity-fitness correlation between individual multilocus heterozygosity (MLH) and an estimate of fitness, body condition. This analysis revealed significant variation in mean MLH and mean body condition among populations but no correlation between them. This result was true even after controlling for non-genetic factors such as sex, season of capture and year of capture. Tests for inbreeding were also negative and suggest snakes are mating randomly within each population. Finally, in Chapter 4, I used stable isotope analysis of carbon and nitrogen to study the patterns of variation in diet over time and space across the range of this species. I employed a set of novel techniques in isotope biology to assess the relative importance of among-population or individual variation and then calculated Manly alpha selectivity indices to determine if prey preference exists for mammalian prey species. Stable isotope analysis revealed diet is variable among populations with shrews and voles being among the most important prey items. A preference for shrews was observed in two of three populations and an avoidance of voles was observed in all three populations.

Overall, my results indicate populations have been isolated for thousands of years and that recent habitat fragmentation caused by humans has had little effect on the patterns of genetic variation in this species. Although populations have been isolated over evolutionary timescales they appear to have retained a large enough size to avoid inbreeding and the negative effects of inbreeding depression. The lack of any correlation between genetics and fitness implies genetics may be less important for the overall fitness
of these snakes when compared to ecological factors. Stable isotopes appear to be a viable method for studying diet in snakes and this analysis represents an important first step in understanding if variation in body condition is a result of variation in diet.
Dedicated to Noni P –

My love for you shall live forever…
Acknowledgements

This work would not have been completed without the valuable help and dedicated assistance of my advisor, H. Lisle Gibbs. I am also very appreciative of help along the way from the remainder of my committee, Peter Curtis and Paul A. Fuerst and from past members of my committee, Thomas Waite and Amanda Rodewald.

There have been many who have helped with collection of data in both the field and laboratory and provided great assistance in completing this work at various stages including, John Adamski, Sarah Baker, Bob Cox, Jennifer D’Agostino, Penny Danielewicz, Jeff Davis, Fred Dierkis, Mike Dreslik, Joanne Earnhardt, Brian Fedorko, Jason Folt, Tony Frazier, James Gibbs, John Harder, Billie Harrison, Benjamin Jellen, Jeff Jundt, Liam Kean, Matt Kowalski, Greg Lipps, Steve Mackessy, Kevin Pangle, Wiline Pangle, Joy Piekiewicz, Melissa Pilgrim, Carrie Pratt, Kent Prior, Kevin Shoemaker, Jessica Wooten, Dan Wylie and Doug Wynn.

Of course, no work would be completed without someone providing the funding and this dissertation is no different. Special thanks to Carolyn Caldwell and the Ohio Department of Natural Resources and Division of Wildlife for not only funding this work but being open to my ideas about massasauga research. All stable isotope work was conducted at
the University of Michigan Biological Station, and Michael Grant donated a generous amount of time and effort to this project. Thanks for sharing all your knowledge. I have also been fortunate to gain an extensive amount of support in teaching while at Ohio State and many have contributed to building upon my teaching foundation and provided me freedom within the classroom to blossom as an instructor. For this I’d like to thank Bob Cox, Cynthia Dassler, Jerry Downhower, Ian Hamilton, John Harder, Wiline Pangle, Joe Raczkowski and Steve Rissing.

Plenty of graduate students over the years have helped with this work but I’d like to specially thank members of the Gibbs’ lab including Christine Anderson, Anthony Fries, Katy Greenwald, Eric McCluskey, Michael Murphy, Kristen Mylecraine, Angelica Poesel, Wayne Rossiter, David Salazar-Valenzuela, Sarah Smiley and many thanks to Jose Diaz for unwavering support in the lab. Graduate students from other labs also need to be thanked: Kody Kuenhl, Mike Sovic, Erin Lindstedt, and Anthony D’Orazio. Over the years I’ve been here there has also been a great deal of support from administrative staff: Cathy Drake, Rene Madsen, Andrea Johnson, Faith Haleem, Gale Azcarraga-Carter, Elisha Starkey, Sue Meier, Chanelle Kinney and Corey Ross.

I especially want to thank my family, most notably my parents and grandparents for their generous support. You taught me nearly everything I know and made me who I am today. And very special thanks to my wife, Missy – What can I say? You have supported me through and through in every way possible without once asking the
question – “Are you almost finished with your dissertation?” You mean more to me than you’ll probably ever realize.
Vita

December 2001………………………………………………………………B.S. Biological Sciences, East Stroudsburg University
2001-2004……………………………………………………………………..M.S. Biological Sciences, East Stroudsburg University
2004-2011………………………………………………………………………..Graduate Teaching Associate, The Ohio State University

Publications


Fields of Study

Major Field: Evolution, Ecology and Organismal Biology
Table of Contents

Abstract .................................................................................................................................................. ii

Acknowledgements ............................................................................................................................. vi

Vita........................................................................................................................................................ ix

List of Tables ......................................................................................................................................... xiv

List of Figures ......................................................................................................................................... xvi

Chapter 1: Introduction ....................................................................................................................... 1
  References............................................................................................................................................. 7

Chapter 2: Similarity of Contemporary and Historical Gene Flow Among Highly-
Fragmented Populations of an Endangered Rattlesnake .................................................................. 10
  Abstract ............................................................................................................................................... 10
  Introduction ......................................................................................................................................... 11
  Materials and Methods ....................................................................................................................... 14
    Sample Localities and Microsatellite Genotyping ........................................................................ 14
    Genetic Variation and Population Differentiation ....................................................................... 14
    Estimating Contemporary and Historical Gene Flow .................................................................. 16
    Genetic Signatures of Changes in Population Size ....................................................................... 19
    Tests of Population History .......................................................................................................... 20
  Results ................................................................................................................................................ 21
    Genetic Data ..................................................................................................................................... 21
    Population Differentiation .............................................................................................................. 22
    Population Variation ...................................................................................................................... 23
    Contemporary and Historical Migration Rates ............................................................................. 24
    Changes in Population Size ............................................................................................................ 26
    Tests of Population History .......................................................................................................... 27
  Discussion .......................................................................................................................................... 27
    Contemporary and Historical Gene Flow ....................................................................................... 28
    High Levels of Population Differentiation .................................................................................... 30
    Variation in Effective Population Size .......................................................................................... 31
  References ............................................................................................................................................ 35
Chapter 3: Lack of a heterozygosity-fitness correlation in isolated populations of a threatened rattlesnake, the eastern massasauga (Sistrurus c. catenatus) ........................................ 51
Abstract .................................................................................................................. 51
Introduction ............................................................................................................ 52
Methods ................................................................................................................... 55
  Data collection and Estimates of Body Condition .............................................. 55
  Estimating genetic diversity .............................................................................. 58
  Genetic Fitness Correlations ............................................................................. 58
  Estimating Inbreeding ....................................................................................... 59
Body Condition Estimates from Zoo Populations ............................................. 60
Results .................................................................................................................... 61
Discussion .............................................................................................................. 62
  No evidence of a genetic-fitness correlation ................................................. 63
  Geographic variation in body condition .......................................................... 65
  Temporal Variation and Fluctuations in Body Condition ............................... 66
Spatial Variation in Ecological Parameters – Differences in Available Resources. 67
Spatial Variation in Ecological Parameters – Differences in Density-Dependent
Factors .................................................................................................................. 68
Implications of Variation in Body Condition ..................................................... 69
Limitations of the study ...................................................................................... 69
Body Condition of Captive Animals in Zoos ....................................................... 71
Conclusions .......................................................................................................... 72
References ............................................................................................................. 74

Chapter 4: Individual and Population Variation in Diet of Massasauga Rattlesnakes
(Sistrurus catenatus) Inferred from Stable Isotope Analysis .................................. 85
Abstract .................................................................................................................. 85
Introduction ............................................................................................................ 87
Methods ................................................................................................................... 91
  Sample Collection ............................................................................................. 91
  Sample Preparation and Analysis .................................................................. 92
  Estimating diet from stable isotope data ......................................................... 93
Hierarchical Structure of Isotopic Variation ....................................................... 95
Single population estimates of diet ................................................................. 96
Prey Preference .................................................................................................... 97
Results .................................................................................................................... 99
  Snake and Prey Isotope Values ..................................................................... 99
  Hierarchical Structure of Isotopic Variation ............................................... 100
  Single population estimates of diet ............................................................. 101
Mammal Abundance and Prey Preference ..................................................... 102
Discussion .......................................................................................................... 102
  SIA versus Conventional Methods of Estimating Diet .............................. 103
Patterns of Variation in Diet – Geographic Variation .................................. 104
Patterns of Variation in Diet – Individual Variation .................................. 106
Patterns of Variation in Diet – Temporal Variation................................. 107
Prey Preference ...................................................................................... 108
Analytical Issues .................................................................................... 110
References.................................................................................................. 112

Bibliography .............................................................................................. 125

Appendix..................................................................................................... 140
Table 2.1. Sample size and genetic diversity estimates for each locality (ID) across 19 microsatellite loci. Observed heterozygosity ($H_o$) and expected heterozygosity ($H_e$) was calculated using FSTAT version 2.9.3. Allelic richness ($AR$) was calculated using MSA based on a minimum of six individuals. The mode and 90% highest probability density limits of $F$ as calculated with 2mod are shown……………………………………...42

Table 2.2. Summary of results from BOTTLENECK for 16 populations for the Wilcoxon’s test for both the two-phase model (TPM) and stepwise model (SMM) along with results from the mode-shift test. Under the mode-shift test, an L-shaped distribution of alleles is expected in the absence of a bottleneck whereas a distribution with a shifted mode is expected in a population that has gone through a bottleneck…………..................................................43

Table 3.1. Results from the general linear mixed effects model with body condition as the response variable and heterozygosity, sex, season of capture and year of capture as the independent variables. Population of capture was included as a random effect and its significance was determined using a likelihood ratio test by comparing models with and without the random effect included……79

Table 3.2. Summary of tests for inbreeding (MLH) calculated as the proportion of heterozygous loci per individual and averaged for individuals within a given population. 95% CI’s were constructed by randomizing alleles for each locus to create 1,000 permuted populations from which mean MLH was calculated for each. P-values were calculated as the number of permutations less than the observed mean value for each population. Significant values after correction for multiple tests are in bold………………………….80
Table 4.1. Summary of results from the Bayesian hierarchical modeling approach modified from Semmens et al. (2009) showing the model numbers and specific terms included in each model. A value of “N” indicates the term was not included in a given model and a value of “Y” indicates the term was included in the model. DIC is the deviance information criterion of model selection. Smaller numbers represent better overall model fit and the box highlights the best model……………………………………...……118

Table 4.2. Results of Manly’s alpha indices calculated for each population and each prey type. Values greater than 0.33 suggest a preference for a given prey type whereas values less than 0.33 suggest avoidance of a given prey type. Values in parentheses represent the 95% confidence interval from 10,000 Manly values randomly generated in R (R Development Core Team 2010)……………………………………119

Table A.1. Summary of 19 polymorphic microsatellite loci. Null alleles were estimated using FreeNa. * represents F_{IS} values significant after correction using the B-Y method (Narum 2006). Number of alleles was calculated from 388 individuals………………………..140

Table A.2. Genetic estimates of pairwise F_{ST} between pairs of populations calculated using MSA based on 10,000 permutations. Values with symbols represent non-significant comparisons after Bonferroni correction…………………………………………………………..141

Table A.3. Summary of data collected across years and season for each population. The number of males and females captured from each population is also shown. The total for each column is shown at the bottom……………………………………………………………………142

Table A.4. List of zoos contributing data for this study with corresponding number of samples from each captive population. Animals in this table were combined to form the “zoo population” for comparison of wild snakes………………………………………………143

Table A.5. A) Number of individual prey items and rattlesnakes in each of the five eastern populations. B) Number of prey items and rattlesnakes from the western population in Colorado………………144
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1.  A) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. Population codes for the sampling locations are found in Table 2.1. Populations with asterisks are those used in the range-wide analysis using STRUCTURE. KL represents both KL-1 and KL-2. State abbreviations are as follows: IL (Illinois); IN (Indiana); KY (Kentucky); OH (Ohio); WV (West Virginia); PA (Pennsylvania); NY (New York); MI (Michigan); ON (Ontario, Canada). B) Inset showing the relative location of the sampling localities within the United States and Canada. 44</td>
<td></td>
</tr>
<tr>
<td>Figure 2.2. Bar plot for STRUCTURE at K = 11 (ln P (D) = -14834) showing the association among genetic clusters and sample locations. Abbreviations for sampling locations are as follows: SSSP (South Shore State Park); PRF (Prairie Road Fen); KL-1 (Killdeer Plains Wildlife Area); WM (Willard Marsh Wildlife Area); GRL-1 (Grand River Lowlands 1); WAIN (Wainfleet Bog); BPNP (Bruce Peninsula National Park); GBINP (Georgian Bay Islands National Park); KPP (Killbear Provincial Park); BS (Bergen Swamp); CS (Cicero Swamp). 45</td>
<td></td>
</tr>
<tr>
<td>Figure 2.3. Bar plot for STRUCTURE at K = 3 showing the association among genetic clusters and sample locations for northeastern Ohio. Abbreviations for sampling locations are as follows: GRL-1 (Grand River Lowlands 1); GRL-2 (Grand River Lowlands 2); GRL-3 (Grand River Lowlands 3). 46</td>
<td></td>
</tr>
<tr>
<td>Figure 2.4 Population estimates of theta ($\theta_N$) from MIGRATE. Values are shown in decreasing order from left to right. Error bars represent the 95% confidence intervals for each estimate. KL represents KL-1. 47</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.5. Likelihood estimates of long-term gene flow among populations at a) range-wide and b) regional scales. Circle sizes reflect relative population size. Values inside the circles are $\theta_N$, shown with 95% confidence intervals. Bold values indicate migration was asymmetrical as determined from non-overlapping 95% confidence intervals and the arrows represent the direction of gene flow among sites. All pairwise estimates of $M (m/\mu)$ are shown with 95% confidence intervals. Black bars represent the position of the Great Lakes in relation to each population.

Figure 2.6. Distribution of differences between contemporary migration rates ($m_c$) estimated using BAYESASS and historical migration rates ($m_h$) estimated using MIGRATE for specific population comparisons ($m_c - m_h$) as shown in Figure 2.4.

Figure 3.1. A) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. Population codes for the sampling locations are as follows: EHSP (Eldon Hazlet State Park); SSSP (South Shore State Park); PRF (Prairie Road Fen); KL (Killdeer Plains Wildlife Area); WM (Willard Marsh Wildlife Area); GRL-1 (Grand River Lowlands 1); GRL-2 (Grand River Lowlands 2); GRL-3 (Grand River Lowlands 3); PA (Western Pennsylvania); BS (Bergen Swamp); CS (Cicero Swamp); WAIN (Wainfleet Bog); BPNP (Bruce Peninsula National Park); KPP (Killbear Provincial Park). B) Inset showing the relative position of the sampling locations within the United States and Canada.

Figure 3.2. Linear relationship between log-transformed mass and log-transformed SVL for all individuals from wild populations as determined by a simple linear regression. These values were then used in a type II regression analysis (standard major axis regression) to calculate scaled mass index as described above. The dashed line represents the mean value (175.5 grams).

Figure 3.3. Mean body condition among wild populations and zoo samples of Sistrurus catenatus. The dashed line shows the mean scaled mass index of 175.17 grams. When using this value as the baseline for body condition, two populations, BPNP and GRL-2, are in above average condition and three populations are in below average condition (EHSP, SSSP and KPP) based on non-overlapping 95% CI (black bars) with the mean.

Figure 3.4 Mean multilocus heterozygosity (MLH) among all populations.
of *Sistrurus catenatus*. Bars represent 95% confidence intervals. Populations are arranged by geographic region within the range of this species from Southwest to Northeast for easier comparison of geographically close populations.................................84

Figure 4.1. A) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. State abbreviations are as follows: IL (Illinois); IN (Indiana); KY (Kentucky); OH (Ohio); WV (West Virginia); PA (Pennsylvania); NY (New York); MI (Michigan); ON (Ontario, Canada). B) Inset showing the relative location of the sampling locations within the United States and the location of the western massasauga locality in Colorado (CO).................................................................120

Figure 4.2. Mean and standard error of carbon ($\delta^{13}$C) nitrogen ($\delta^{15}$N) isotopic composition prior to any correction for fractionation for each prey taxa and massasauga rattlesnakes. Results are shown for each of five eastern massasauga rattlesnake populations (ILL, KL, PRF, GRL-2, CS) and one western massasauga rattlesnake population (CO). Massasauga mean values are displayed within a box to highlight their position relative to potential prey items.........................121

Figure 4.3. A) Distribution of prey items in the diet of A) eastern massasauga rattlesnakes and B) western massasauga rattlesnakes determined using the fully Bayesian model for each population independently. Each column represents specific prey taxa while each row represents a given population. Median values are reported above each individual graph and the taxa representing the highest proportion of the diet is marked with a box.................................................................122

Figure 4.4. Distribution of prey items in the diet of western massasauga rattlesnakes determined using the fully Bayesian model. Each column represents specific prey taxa. Median values are reported above each individual graph and the taxa representing the highest proportion of the diet is marked with a box.................................................................123

Figure 4.5. Comparison of the abundance of mammalian species as determined by mammal trapping at each site (Piekiewtocz unpubl. data) (upper graph) and the percentage of each mammalian prey type as calculated using SIA under the fully Bayesian model (lower graph) from each of three populations.................................................................124

xviii
Figure A.1. $\Delta K$ calculated with the Evanno method (2005) from the range-wide STRUCTURE analysis. There is strong support for a $K$ of 11 as determined by the relative height of the peak at this value.
Chapter 1: Introduction

The genetic consequences of habitat fragmentation and subsequent isolation of animal populations has garnered a great deal of attention in the past decade. As populations become isolated from each other, this can lead to changes in the levels of genetic diversity in each populations often by altering microevolutionary forces such as gene flow, genetic drift and natural selection at the local level (Keyghobadi 2007). For roughly 80 years, individuals have asked questions about the genetic effects of population isolation and my understanding of the effects of this process has greatly expanded (Nei & Tajima 1981; Wright 1931). Often, habitat fragmentation leads to a reduction in overall population size while increasing the degree of isolation among populations. This often leads a population to experience increased levels of genetic drift, due to random evolutionary changes caused by small population size, and decreased levels of gene flow between populations. Ultimately, drift in combination with low levels of gene flow will lead to a reduction in overall genetic diversity within populations and increase genetic differentiation among populations (Keyghobadi 2007).

Isolated populations are also in danger of further reductions in genetic diversity through bottlenecks (Piry et al. 1999) that can ultimately lead to local extinction of populations.
and habitat fragmentation and isolation can lead to other genetic problems that may further reduce the overall fitness of the population. Inbreeding, the mating of closely related individuals who share common ancestry, can also become a problem in small populations because the likelihood that two individuals are closely related is increased (Keller & Waller 2002). Populations experiencing the effects of inbreeding often further suffer from inbreeding depression, a reduction in individual fitness due to either the expression of deleterious alleles in homozygotes or an increased advantage to heterozygous individuals (Lynch and Walsh 1998). The effects of inbreeding depression are often witnessed in life history (e.g. number and size of offspring), or morphological traits (e.g. body condition) and can lead to drastic reductions in overall population fitness (DeRose & Roff 1999).

Inbreeding and inbreeding depression have both been documented in wild populations and are often linked to small, isolated populations that display reduced levels of genetic variation (Madsen et al. 1996). Inbreeding is measured in many ways in nature but one common method involves comparing observed heterozygosity to estimates of heterozygosity in a randomly mating population, as populations that are experiencing inbreeding are expected to show increased levels of homozygosity and thus lower levels of heterozygosity (Castric et al. 2002). This is also commonly done through analysis of relatedness of individuals in a population under the assumption that a population with higher overall relatedness may be experiencing higher levels of inbreeding (Castric et al. 2002). Making a connection between inbreeding and inbreeding depression often involves correlating individual levels of heterozygosity with some estimate of individual
fitness (Grueber et al. 2008). Often, the fitness traits estimated are life history traits but morphological traits are often measured as well. One common estimate of fitness is to obtain an estimate of body condition as the residuals from a regression of mass and body length (Peig & Green 2009). Then, estimates of individual heterozygosity are regressed against body condition (individual fitness), commonly referred to as genetic-fitness correlations (Chapman et al. 2009). In this method, a significant positive linear relationship would suggest individuals with higher MLH are in better body condition and vice versa indicating the population may be suffering from inbreeding depression.

Though genetic analysis can yield informative results about fragmented and isolated populations, genetic diversity and patterns are not the only important factors for fragmented populations. In isolated populations of endangered species it becomes especially important to understand variation in significant ecological components that might directly affect an individual’s fitness as well. One such ecological component is diet. Diet is a key component to the ecology of a species and gaining insights into species-wide diet variation can help answer questions about a species foraging ecology (Godley 1980), predator-prey co-evolution (Williams et al. 2004) and the evolution of specialized morphological structures (Greene 1983).

Studying a species diet is not always easy though. Many species live secretive lives and are rarely observed, let alone viewed while feeding. Similarly, many species feed at night or in extreme habitats that make observing an organism’s feeding habits difficult. Because of this, there are many species for which this basic biological information is
missing or incomplete. These difficulties in determining diet have lead to biologists on techniques that provide an estimate of diet without directly observing organisms feeding (McKechnie 2004). Stable isotope analysis has been used ever increasingly in ecological studies in an attempt to study diet (Gannes et al. 1997). This technique is based on the fact that in nature many elements have at least two stable isotopes, one of which is much more common than the other (Hobson & Wassenaar 1999). This technique essentially measures the relative amount of each isotope for a particular element thereby giving an isotope ratio (e.g. $C^{13}:C^{12}$). Elements most commonly used in ecological studies include hydrogen ($H^2:H^1$), carbon ($C^{13}:C^{12}$) and nitrogen ($N^{15}:N^{14}$) (Gannes et al. 1997). These elements have been shown to vary based on geography (Pilgrim 2005), trophic level (Urton & Hobson 2005) and vegetation type (Koch et al. 1995) and as a result, different potential prey species may exhibit different isotopic signatures if they are feeding at different trophic levels or on different types of prey themselves. When a predator feeds on prey, the isotope signature of the prey is incorporated into the predator’s tissue and as such, by measuring the ratios of these isotopes in various prey species as well as the predator, one can estimate the components of different prey in the diet of the predator (Hobson & Wassenaar 1999).

Eastern massasauga rattlesnakes (Sistrurus c. catenatus) are small to medium-sized vipers currently listed as a subspecies of the massasauga rattlesnake (Sistrurus catenatus) (Campbell & Lamar 2004). This species was historically thought to occur in a series of rather connected populations throughout the Midwestern United States stretching into southern Ontario, Canada but currently, eastern massasaugas occur in highly isolated
populations thought to be a result of habitat fragmentation due to human alteration of the landscape. This combined with perceived recent declines in population size caused the United States Fish and Wildlife Service to categorize the eastern massasauga as a candidate sub-species for endangered status (Szymanski 1998). In combination, the above issues of habitat fragmentation and population isolation make the eastern massasauga an ideal study species to compare the effects of historical and contemporary gene flow and further investigate the effects of inbreeding and inbreeding depression in these populations. Similarly, due to this species’ endangered status and secretive nature, little is known about massasauga diet, particularly patterns of diet variation across the species range. In the following paragraphs I highlight each chapter of my dissertation and provide a brief summary of my work addressing some of the above concerns of isolated populations using data collected from a comprehensive sample of populations across the range of the eastern massasauga.

In Chapter 1, I conduct a comprehensive analysis of the genetic characteristics of eastern massasauga rattlesnake populations using data from 19 microsatellite loci. I had three major goals with this chapter: 1) Highlight the overall patterns of genetic diversity on a regional and range-wide scale; 2) Investigate each population for signs of population size reduction and geographic structure; 3) Compare historical and contemporary estimates of gene flow (migration rates) among populations. Overall, I find populations are highly differentiated and each population represents a unique genetic cluster. There is no evidence for any population size reduction suggesting populations have been stable over evolutionary timescales but that populations are highly variable in effective population
size. Historical and contemporary migration rates are low and similar suggesting populations have been isolated for thousands of years suggesting that recent habitat fragmentation due to human alteration of the landscape is not responsible for creating the genetic patterns observed in this species.

In Chapter 2, I use an estimate of multilocus heterozygosity (MLH) from 19 microsatellite loci to investigate the potential fitness costs of massasaugas living in fragmented populations isolated for thousands of years. First, I test for any evidence of inbreeding by comparing observed estimates of MLH with estimates of MLH in a randomly mating population. Next, I test for evidence of inbreeding depression by regressing MLH against my estimate of fitness, body condition. Finally, to establish a benchmark for body condition in wild snakes, I compare individual estimates of body condition in wild snakes with those of captive snakes housed in zoos. I find no evidence of inbreeding or inbreeding depression in these populations; however, body condition is quite variable among populations and cannot be explained by genetics. I also show that captive animals housed in zoos provide a good benchmark for body condition in wild snakes with some wild populations in good body condition while others are in relatively poor body condition.

Finally, in Chapter 3, I use stable isotope analysis of carbon and nitrogen as a tool to study the patterns of variation in diet in S. catenatus. I was particularly interested in investigating the pattern of intraspecific variation in diet across the range using a set of novel methods developed specifically for stable isotope analysis. First, I use a modeling
technique to assess the importance of within and among population variation in diet.

Next, I use Bayesian mixing models to estimate the contribution of each potential prey item to the diet of massasauga rattlesnakes. Finally, dietary preference for specific mammalian prey was examined in these populations. Overall, I found population-level variation in diet with no evidence that diet varies among individuals within each population. Shrews comprised the bulk of the diet in three out of five populations. Massasaugas displayed a preference for shrews in two of three populations and an avoidance of voles in all three populations. These findings are then discussed in terms of their evolutionary and ecological implications.

References


Chapter 2: Similarity of Contemporary and Historical Gene Flow Among Highly-Fragmented Populations of an Endangered Rattlesnake

Abstract

Populations of endangered taxa in recently fragmented habitats often show high levels of genetic structure, but the role that contemporary versus historical processes play in generating this pattern is unclear. The eastern massasauga rattlesnake (Sistrurus c. catenatus) is an endangered snake that presently occurs throughout central and eastern North America in a series of populations that are isolated due to habitat fragmentation and destruction. Here, I use data from 19 species-specific microsatellite DNA loci to assess levels of genetic differentiation, genetic effective population size, and contemporary and historical levels of gene flow for 19 populations sampled across the range of this snake. Eastern massasaugas display high levels of genetic differentiation (overall $F_{st} = 0.21$) and a Bayesian clustering method indicates that each population represents a unique genetic cluster even at regional spatial scales. There is a two-fold variation in genetically effective population sizes but little genetic evidence that populations have undergone recent or historical declines in size. Finally, both contemporary and historical migration rates among populations were low and similar in magnitude even for populations located less than 7 km apart. A test of alternate models of
population history strongly favors a model of long-term drift-migration equilibrium over a recent isolation drift-only model. These results suggest that recent habitat fragmentation has had little effect on the genetic characteristics of these snakes, but rather that this species has historically existed in small isolated populations that may be resistant to the long term negative effects of inbreeding.

Introduction

Habitat loss and fragmentation can alter the levels and distribution of genetic diversity among populations in landscapes heavily impacted by humans (Keyghobadi 2007). Low genetic diversity can reduce population viability via genetic mechanisms such as inbreeding depression, loss of adaptive potential, and the accumulation of deleterious mutations (Frankham et al. 2002). Empirical studies show high levels of genetic structure and variable population sizes of organisms in fragmented habitats (Hartl et al. 2005; Martinez-Solano et al. 2005; Redeker et al. 2006; Schwartz & Karl 2005). However, such studies cannot infer a causal relationship between fragmentation and genetic characteristics because they often lack comparison with genetic characteristics of populations in unfragmented habitats (Keyghobadi 2007). In fact, evolutionary processes operating over historical timescales (pre-dating the onset of human impacts) could be responsible for the observed genetic structure, and the observation of structure in contemporary populations in fragmented habitats may be coincidental rather than due to the direct effects of fragmentation (Cunningham & Moritz 1998; Hansen et al. 2009; Pavlacky Jr et al. 2009; Vandergast et al. 2007).
It is especially difficult to disentangle the relative role of contemporary and historical processes when most of the range occupied by an endangered taxon has been heavily affected by humans, making comparisons of populations in fragmented and unfragmented habitats impossible. Various approaches have been used to address this problem, including comparing patterns of differentiation of molecular markers with differing mutation rates, (Johnson et al. 2003) and the use of historical landscape data or historical specimens (Martinez-Cruz et al. 2007; Pertoldi et al. 2007). Here, I adopt another approach, that of analyzing a dataset with two analytical techniques that extract information on evolutionary processes operating at distinct time scales (e.g. (Hansen et al. 2009). These time scales bracket the period during which anthropogenic effects began to affect the landscape. Specifically, I compare assignment tests versus coalescent techniques to estimate migration rates over few versus many generations, respectively (Beerli & Felsenstein 2001; Wilson & Rannala 2003). In addition, I use measures of genetic variation to detect changes in population size over different time scales (Cornuet & Luikart 1996). These complimentary approaches allow us to assess the impact of historical and contemporary processes on the genetic characteristics of an endangered North American snake, the eastern massasauga rattlesnake (*Sistrurus c. catenatus*).

Eastern massasaugas are small to medium size vipers that are currently classified as a subspecies of the massasauga rattlesnake (*S. catenatus*) although recent phylogenetic analyses using variation in nuclear and mitochondrial DNA loci suggest that this subspecies is a highly distinct taxon deserving of species status (Kubatko et al. 2011).
These snakes represent an ideal system in which to compare the effects of current and historical gene flow. Historically, massasaugas were thought to occur throughout the Midwestern United States and Canada as a series of connected populations in what is known as the Pleistocene Prairie corridor (Cook 1992). However, many extant massasauga populations are currently found in relatively small, isolated patches of habitat surrounded by heavily modified landscapes (Szymanski 1998). Population declines throughout the range attributed to habitat fragmentation and destruction have lead the United States Fish and Wildlife Service to list this snake as a candidate sub-species for endangered or threatened status within the United States (Szymanski 1998) and it is also classified as a federal Species at Risk in Canada (Canada 2009).

Here, I conduct the most comprehensive analysis of the genetic characteristics of eastern massasauga populations to date using data from 19 microsatellite loci from 19 sampling localities present at both range-wide and regional scales. Population genetic analyses seem most appropriate to examine spatial patterns of genetic diversity as there is little evidence of phylogeographic structure within this taxon although the analysis are limited (Kubatko et al. 2011). My specific goals were to: 1) Characterize the overall patterns of genetic diversity as well as regional and range-wide genetic structure of eastern massasauga populations; 2) Look for genetic signatures of population size changes; 3) Investigate the levels of recent and historical gene flow among populations; 4) Test among alternate models of demographic history namely migration-drift equilibrium or fragmentation followed by drift with no gene flow (Ciofi et al. 1999). I aim to disentangle the effects of short versus long-term processes on patterns of genetic
variation. This will allow us to assess the effects of anthropogenic habitat modification on this taxon, thereby informing conservation efforts towards this snake.

Materials and Methods

Sample Localities and Microsatellite Genotyping

I collected samples from adult massasauga rattlesnakes (N = 388) from 19 localities (n= 6 – 68 samples per site) throughout their range (Table 2.1, Figure 2.1). Whole genomic DNA was extracted from 100 μL of blood using phenol-chloroform and samples were genotyped at 19 polymorphic microsatellite loci using primers and protocols described by Anderson et al. (2010). Briefly, 3-4 loci were simultaneously amplified using the Qiagen multiplex PCR kit in 10 μl reactions containing 5 μl of Qiagen multiplex PCR Master Mix, 1 μl of primer mix (containing 2μM of each primer), and 3.5 μl of Rnase-free water (see Anderson et al. [2010] for details of locus-specific PCR conditions). Amplified products were run on an ABI 3100 genetic analyzer and allele sizes scored using GENEMAPPER 3.7 software.

Genetic Variation and Population Differentiation

I used FSTAT v2.9.3 (Goudet 2001) to determine the number of alleles, observed (H₀) and expected heterozygosity (Hₑ), and F₁S for each locus. For each population, MSA v4.05 (Dieringer & Schlüterer 2003) was used to calculate allelic richness (AR) on a
minimum of 6 individuals. Departures from Hardy-Weinberg equilibrium for each locus and population were assessed using GENEPOP v1.2 (Raymond & Rousset 1995) with the critical $P$-value corrected for multiple tests using the Benjamini and Yekutieli (B-Y) method (Narum 2006). I then used FreeNA (Chapuis & Estoup 2007) to estimate a null allele frequency at each locus in each population.

I used two approaches to assess genetic differentiation among sample locations. First, I used MSA v4.05 to estimate and test the significance of Weir and Cockerham’s (1984) unbiased $F_{ST}$ estimator across all populations as well as for all pair-wise combinations of populations. Second, I used the Bayesian clustering analysis implemented in STRUCTURE v2.3.1 (Pritchard et al. 2000) to infer the number of distinct genetic groups observed in my data. Because of marked heterogeneity in the distances between my sampling localities (Fig. 2.1) I conducted STRUCTURE analyses at two different spatial scales. First, I conducted a large-scale range-wide analysis among populations separated by more than 50 km (see Figure 2.1). For regional groups with multiple populations within 50 km of each other (NE Ohio, central Ohio, western Pennsylvania and southern Illinois), I randomly chose one population as “representative” of that specific region to use in the range-wide analysis. This resulted in an initial data set consisting of 13 populations. Small sample size has been shown to influence effectiveness of STRUCTURE in assigning individuals (Evanno et al. 2005) and so I next removed populations containing fewer than 10 individuals from the data set. This resulted in a second range-wide dataset of 11 populations that was analyzed as described below. Second, I also conducted separate regional analyses among each of 4 sets of populations
located less than 50 km from each other (NE Ohio (n=3), central Ohio (n=2), western Pennsylvania (n=3) and southern Illinois (n=2)).

For both sets of analyses, each STRUCTURE run consisted of a burn-in of 50,000 MCMC followed by 300,000 iterations using the admixture model with sampling localities as priors and correlated allele frequencies. The plot of (ln Pr(X|K) and the likelihood of the runs were used to determine if the runs converged by observing whether the values converged to a particular likelihood with no significant deviations. I performed five runs for each value of K ranging from 1 to 15 (range-wide scale) or 1 to 5 (regional scale). To determine the most likely value of K, I used the Evanno et al. (2005) method implemented in the program STRUCTURE HARVESTER (Earl 2009) which determines the second order rate of change in the distribution of L (K). I used DISTRUCT v1.1 (Rosenberg 2004) to display the results from the best run among the five runs for the most likely value of K for both the range-wide and regional analyses.

Estimating Contemporary and Historical Gene Flow

To compare migration rates over contemporary and historical timescales, I used the programs BAYESASS (Wilson & Rannala 2003) and MIGRATE (Beerli 2002) respectively. Both programs generate parameters from which a comparable measure of gene flow can be inferred (m: proportion of population consisting of genetic migrants per generation) but each estimates this parameter over a different time scale: BAYESASS uses a Bayesian approach and MCMC sampling to generate m values which reflect migration rates over “…the last few generations” (Wilson & Rannala 2003) which I take
to be < 5 generations or 15 – 30 years given variation in the estimated generation time (3 - 6 years) of these snakes (Miller 2005). In contrast, MIGRATE uses the coalescent to jointly estimate the relative effective population size $\theta_{Ne} (4N_e \mu)$ and asymmetric gene flow $M (m/\mu)$ between pairs of populations over much longer (~4N_e generations in the past; Beerli 2008) periods of time (~1000’s of years). As such, these migration rates provide estimates of gene flow that post- and predate the estimated time (~300 yr bp: (Monkkonen & Welsh 1994)) when humans began to significantly alter the habitats in which these snakes currently live.

For BAYESASS analyses I used all 19 populations in each run but only report the values that mirror the comparisons carried out in MIGRATE (see below). Initially, each run consisted of 3x10^6 iterations with the chain sampled every 2000 iterations. A burn-in of 10^6 was used and delta values were adjusted to ensure that 40-60% of the total changes were accepted. As recommended by Faubet et al. (2007) I performed 10 runs of BAYESASS each with a different initial seed and then used a Bayesian deviance measure to determine the run which best fit the data (Spiegelhalter 2002). For this best fit run, I then ran the analysis again using the seed from the best fit run but increased the run length to 3 x10^7 iterations. The results presented are from this final run.

I used MIGRATE to estimate long-term estimates of gene flow and effective population size. As with the STRUCTURE analyses, I conducted the MIGRATE analyses at both regional and range-wide spatial scales. At the regional scale I estimated gene flow between clusters of discrete sample locations all within 25 km of each other (see above).
At the range-wide scale, I did not estimate migration rates for every possible population combination (Palstra et al. 2007). Rather, due to the limited gene flow suggested by high levels of differentiation (see below and (Gibbs et al. 1997) I analyzed a modified stepping-stone model of migration using the following criteria to determine the sets of populations for which migration rates were estimated. If two populations could be connected with a straight line not passing through another population, then migration rates were estimated for these populations. The Great Lakes were viewed as barriers to gene flow and therefore migration rates were not estimated for populations separated by one or more of the great lakes (except for BPNP and GBINP due to their close geographic proximity). This model was used as input for the migration model option in MIGRATE. For each of the regional sets of populations, I randomly chose one of the sampling localities to use in range-wide MIGRATE analyses.

For both the range-wide and regional analyses, MIGRATE was run a total of three times. The first two runs were shorter (10 short chains of 10,000 sampled, 500 recorded and 3 final chains of 100,000, 5000 recorded) and used to verify the MCMC was estimating the parameters correctly. Then, a final run (10 short chains of 10,000 sampled, 500 recorded and 3 final chains of 500,000 sampled and 25,000 recorded) was performed and $M$ values from this final run are reported here. The initial run used an estimate of $F_{ST}$ as a starting parameter to calculate $\theta_{Ne}$ and $M$ and each subsequent run used the ML estimates from the previous run as new starting parameters. Four-chain heating at temperatures of 1, 1.5, 3 and 10,000 was implemented to increase the efficiency of the MCMC. Because parameter estimates from the final run were similar to the results from the shorter runs I
assumed that the final run had converged and I present results from this final run. For all analyses, only 20 randomly-chosen individuals were analyzed per population because of computational demands and evidence that greater than 20 individuals does not increase the accuracy of parameter estimates in MIGRATE analyses (Kuhner 2006).

To compare migration rates generated by BAYESSASS and MIGRATE, I used the values of \( m \) directly generated by BAYESSASS and estimated \( m \) from values of \( M (m/\mu) \) generated by MIGRATE by dividing all \( M \) values by an estimated mutation rate of \( 5\times10^{-4} \) (Garza & Williamson 2001). For this comparison, I excluded all estimates of migration rates associated with OJIB due to the inflated contemporary migration rates linked to this population (see below). To statistically test for similarity in migration rates, I performed a Mantel test using ztmantel (Bonnet & Van de Peer 2002) with 10,000 permutations to test for an association between contemporary and historical values of \( m \).

**Genetic Signatures of Changes in Population Size**

I used two methods implemented in BOTTLENECK to detect genetic bottlenecks. First, I used the Wilcoxon's sign rank test, which examines whether populations exhibit a greater level of heterozygosity than predicted in a population at mutation-drift equilibrium. This test is most sensitive at detecting bottlenecks occurring over approximately the last 2-4 \( N_e \) generations. Second, I used the mode-shift test which is most appropriate for detecting population declines which have occurred more recently, specifically over the last few dozen generations (Cornuet & Luikart 1996; Luikart *et al.* 1998). This test is based on
the idea that non-bottlenecked populations at mutation-drift equilibrium are expected to have a large proportion of alleles at low frequency and a smaller proportion of alleles at intermediate frequencies (L-shape distribution) (Luikart et al. 1998). In bottlenecked populations, the distribution will shift towards one in which a smaller proportion of alleles are found at low frequency (< 10%) than at intermediate frequency because the alleles at low frequency are those most likely lost during a bottleneck (shifted mode distribution). I performed 10,000 simulations using the program BOTTLENECK 1.2.02 (Piry et al. 1999) in 16 populations (excluding VEN, SGL-95 and KL-2) under the stepwise mutation model (SMM) and the two-phase model (TPM) with 95% single step mutations and 5% multi-step mutations and a variance of 12 as recommended by Piry et al. (1999). P-values from the Wilcoxon’s test were used as evidence for bottlenecks occurring at each time scale and were assessed for significance at the 0.05 level.

Tests of Population History

Finally, I evaluated which of two models of population history (gene flow versus drift only) best describes the processes leading to the current population structure of eastern massasaugas using 2mod, a coalescent model that compares the likelihoods of each model using MCMC simulation (Ciofi et al 1999). The gene flow model assumes populations are at drift-migration equilibrium and uses the gene frequencies within populations to estimate the relative strength of drift versus gene flow for each population. The drift model assumes a historical panmictic population separated into many smaller populations that have since been evolving independently through drift alone in the
absence of gene flow. Both models assume mutation is small and calculate a parameter $F$, the probability that two alleles in a given population share a common ancestor. I ran the program twice for each model with 100,000 MCMC updates; the first 10% of the output was discarded as burnin. Results from the two runs were combined and the probability of each model was calculated as the number of draws for a given model out of the total draws. I used Bayes factor to describe the probability of the most likely model over the probability of the other model. I also report the mode and 90% Highest Posterior Density (HPD) limits for the posterior distribution of $F$ for each population using the function locfit in R (Loader 1999).

Results

Genetic Data

I genotyped 388 eastern massasaugas from 19 localities at 19 polymorphic microsatellite loci (Table 2.1). A total of 262 alleles were observed with numbers of alleles per locus ranging from 4 (Scu 206) to 29 (Scu 200 and Scu 216) with an average of 13.8 alleles per locus. Observed heterozygosity ($H_O$) among loci was variable ranging from 0.16 (Scu 206) to 0.76 (Scu 212) with an average of 0.58 over all samples (Table A.1).

Seven of the nineteen loci were significantly out of HWE after adjusting the critical p-value using the B-Y method (Benjamini & Yekutieli 2001; Narum 2006) (Table A.1). However, of the 361 tests performed across all loci by all populations, only 22 (6%) were significant after correction with the B-Y method. Loci not in HW appear to be random
with respect to population. Only locus Scu 215 showed significant deficits in more than two populations (6 total) and two loci (Scu 216 and Scu 217) each had significant deficits in only two populations. Estimates of the frequency of null alleles using FreeNa were \( \leq 8\% \) for all loci.

*Population Differentiation*

Overall, 21% of the genetic variation present in the dataset was due to among locality variation (\( \Theta_{Fst} = 0.206, 95\% \text{ CI: } 0.18-0.23, p < 0.0001 \)). Pair-wise \( \Theta_{Fst} \) ranged from 0.08 between BPNP and KL-1 to 0.38 between GRL-2 and PRF. For all pairwise comparisons (except KL-1 to KL-2), \( q_{Fst} \) was significant after applying the B-Y correction for multiple tests (Table A.2) and there was no evidence of isolation by distance (\( r = 0.22, p = 0.15 \)).

At a range-wide spatial scale, initial analyses using STRUCTURE using 13 populations (see above) identified genetic clusters that closely corresponded to the sampling localities with the exception of the two localities with the smallest sample sizes (\( n \leq 9 \)) (OJIB and JEEC). Individuals from these localities were often assigned to clusters containing individuals from other sampling localities and the probability of assignment for individuals was often low (< 50%). As mentioned above, small sample size has been shown to reduce effectiveness of STRUCTURE in assigning individuals (Evanno *et al.*, 2005) and so I assumed that the low assignment probabilities for these individuals were an artifact of low sample size and removed these populations from the analysis. With the modified dataset containing 11 sampling localities, the Evanno method, determined the
most likely value of $K$ was 11 ($\Delta K_{\text{MAX}} = 25.15$) (Figure A.1). Of the 5 runs for $K = 11$, the run with the highest posterior probability ($\ln \Pr(X|K) = -14834.2$) had a highly non-random association of individuals corresponding strongly to sampling localities (Figure 2.2).

STRUCTURE was also used to estimate the number of genetic clusters among sets of regional populations in NE Ohio, central Ohio, western PA and southern Illinois. Other than the two populations 1.5 km apart at Killdeer Plains (KL-1 and KL-2) where $K = 1$ was favored, all regional analyses showed strong evidence for separate genetic clusters that correspond tightly to the sampling localities. Specifically, for the two sites in Illinois, STRUCTURE found overwhelming evidence for a $K = 2$ ($\Delta K_{\text{MAX}} = 130.1$) and individuals were assigned to clusters corresponding to their individual sampling localities. In NE Ohio, analyses of three sample sites gave a $K = 3$ ($\Delta K_{\text{MAX}} = 441.4$) and the three clusters were again defined by sampling locality (Figure 2.3). Finally, for the three populations in western Pennsylvania, a $K = 3$ received the greatest support ($\Delta K_{\text{MAX}} = 288.8$) and once more the three clusters tightly correspond to the three sampling localities.

Population Variation

At a regional level, the two populations from New York exhibited the lowest level of heterozygosity (BS: $H_E = 0.49$; CS: $H_E = 0.52$). Sites in northeast Ohio and western Pennsylvania displayed low to moderate levels of genetic diversity as a group. Southern
Illinois along with populations in southern Ontario in general showed moderate to high levels of genetic diversity and allelic richness. Genetic diversity in central and western Ohio was quite variable. This region contained populations with the highest levels of genetic diversity (KL-1, KL-2 and WM) while PRF was among the least diverse populations in the study ($H_E = 0.56$). Estimates of $\theta_{Ne}$ (based on MIGRATE analysis described below) ranged from 0.16 (WAIN) to 2.20 (BPNP) representing an order of magnitude difference in the long-term effective size of populations across the range (Figure 2.4).

**Contemporary and Historical Migration Rates**

Multiple runs of BAYESASS yielded consistently low estimates of contemporary gene flow among populations at both range-wide and regional spatial scales. Excluding comparisons with OJIB, only 1 of 171 possible migration rates among sites were greater than 1% (KL-2 into KL-1) and most estimates of migration had 95% confidence intervals that overlapped zero indicating little to no recent migration among populations (results not shown). In central Ohio, KL-1 and KL-2, which are separated by 1.5 km showed some evidence of recent exchange of individuals. This result likely reflects real migration due to the close proximity of these two sites and the small non-significant $q_{Fst}$ values. During preliminary analyses, immigration rates into OJIB were always high, often at 30%, and source populations were variable. For the final run, migration rates into OJIB were all greater than or equal to 0.01 and the greatest contribution came from
WAIN (9.5%) and GRL-3 (3.7%). However, for reasons discussed below, these results were interpreted as due to analytical artifacts.

Estimates of historical migration rates ($M$) calculated using MIGRATE also revealed little to no migration among populations over the long-term at either spatial scale. At the range-wide scale, estimates of $M$, ranged from 0.36-11 (Figure 2.5a). Of the 27 pairwise estimates of $M$ for sets of populations, 10 were significantly asymmetrical (non-overlapping 95% confidence intervals for pairwise migration) and for all 10 the higher value of $M$ was for immigration from the population with the larger theta to the population with the smaller theta. The number of migrants per generation, $N_{e}m$ (product of theta and $M$ divided by four), ranged from 0.01-1.21. My migration matrix resulted in 54 migration parameters among 13 populations and only four translated into $N_{e}m$ estimates greater than 1 (range: 1.06 – 1.21).

Results from each of the three regional analyses also revealed little to no migration as estimated by $M$, even among geographically close populations (Figure 2.5b). Estimates of $\theta_{Ne}$ from both sites in southern Illinois were the same (EHSP mean 1.65, 95% CI 1.57-1.92; SSSP 1.65, 1.59-1.85) and $N_{e}m$ was less than one (range: 0.64-0.79). $\theta_{Ne}$ varied slightly among populations in western Pennsylvania (JEEC 1.30, 1.24-1.36; SGL-95 0.28, 0.27-0.34; VEN 1.51, 1.44-1.73) but all estimates of $N_{e}m$ were less than one (range: 0.05-0.23). Similarly, in northeast Ohio, there was a 4-fold difference in $\theta_{Ne}$ among population (GRL-1 0.84, 0.82-0.93; GRL-2 0.38, 0.37-0.42; GRL-3 0.19, 0.19-0.21). All estimates of $N_{e}m$ from northeast Ohio were also less than one (range: 0.01-0.38).
The distribution of historical migration rates (from MIGRATE, $m_h$ calculated from $M$, see above) subtracted from contemporary migration rates ($m_c$, from BAYESASS) was centered on zero (Figure 2.6). Results from the Mantel test were significant ($P < 0.05$, $r = 0.35$) indicating that the two matrices of contemporary and historical migration values are significantly correlated with each other.

Changes in Population Size

Most populations showed no genetic evidence for changes in population size over either the long or short term. The Wilcoxon test as implemented in BOTTLENECK under the SMM revealed evidence of a past genetic bottleneck in only three out of 16 populations (PRF, $P < 0.001$; GBINP, $P = 0.04$; KPP, $P = 0.04$; uncorrected for multiple tests; only PRF remained significant after B-Y correction of critical value; Table 2.2). Under the TPM, only PRF showed evidence of a bottleneck (PRF, $P < 0.05$; Table 2.2). The mode shift test deviated from the expected L-shaped distribution in 3 out of 16 populations (OJIB, JEEC and GRL-1). However, two of the three significant results come from populations with low sample size (JEEC and OJIB) and the frequency of alleles in the lowest and second lowest size class were similar for both populations suggesting only a very slightly shifted mode and hence little evidence for a recent bottleneck (0.26 and 0.28, OJIB; 0.27 and 0.30, JEEC). There was a much greater difference between alleles at the lowest frequency and those in the second lowest frequency in GRL-1 (0.22 and
0.35) suggesting some evidence of a weak bottleneck over the last few dozen generations (~ 100 – 200 yrs).

**Tests of Population History**

Consistent with the similarity of short and long term estimates of gene flow, a long-term gene flow population model was much more strongly supported than a pure drift model (Prob. of drift model = 0.0, Prob. of gene flow model = 1.0). This indicates that populations have been at drift-migration equilibrium for a substantial period of time. Values for $F$ varied across populations indicating drift operates more strongly in some populations than others. BPNP and KL-1 were the two populations least affected by drift ($F = 0.08$ and 0.07 respectively) while BS and GRL-1 were most affected ($F = 0.41$ and 0.35 respectively) (Table 2.1). Estimates of $\theta_{Ne}$ from the program MIGRATE were correlated with $F$ values from the 2mod analysis which is consistent with the idea that smaller populations have been more strongly influenced by drift (Rank Correlation, $r = -0.34$, $P < 0.05$).

**Discussion**

My major results are that 1) Both contemporary and historical migration rates among populations were low and similar in magnitude even for populations located less than 7 km apart, and a test of alternate models of population history strongly favors a model of long-term drift-migration equilibrium over a recent isolation drift-only model; 2) Geographically discrete massasauga populations exhibit a high degree of genetic
structure even at regional spatial scales; 3) There is a two-fold difference in genetic
effective population sizes but little genetic evidence that populations have undergone
recent or historical declines in size. I discuss the evolutionary and conservation
implications of these findings below.

Contemporary and Historical Gene Flow

My analyses show that historical and contemporary levels of migration between
genetically distinct populations were small and similar in magnitude. Low levels of
contemporary migration are not surprising since most extant populations are found in
habitat islands that are completely surrounded by heavily-modified landscapes (e.g.
agriculture) making dispersal between populations highly unlikely (Szymanski 1998).
More surprising is that historical migration rates are also low, given that these represent
averages over time periods that predate colonization and subsequent landscape
modification by European settlers in North America (Pielou 1991; Schmidt 1938). These
results strongly imply that the high levels of structure currently observed are not a
consequence of living in a highly-fragmented habitat, but rather that extremely limited
dispersal is a long-standing biological feature of this taxon. Studies of the spatial ecology
of massasaugas confirm that limited dispersal and localized daily movements are a
general phenomenon in this species (Johnson 2000; Reinert & Kodrich 1982). This may
be driven by significant advantages to philopatry for as yet unknown biological reasons
(Gibbs et al. 1997; Gibbs & Weatherhead 2001) and/or long-term habitat heterogeneity
that has existed independent of human impacts. My results add to the increasing number
of studies that emphasize the importance of historical factors in molding patterns of genetic variation within and between populations of species that occupy highly impacted landscapes (Cunningham & Moritz 1998; Hansen et al. 2009; Pavlacky Jr et al. 2009; Vandergast et al. 2007).

One evolutionary implication of similar migration rates over time is that these snakes have a long history of living in relatively small isolated populations that have been dominated by the effects of genetic drift. As such, these populations, especially the smaller ones, may have a history of small to moderate levels of inbreeding which could minimize the effects of inbreeding depression due to the repeated exposure of deleterious recessive alleles to selection; although behavioral mechanisms that result in the avoidance of mating with kin thereby minimizing inbreeding may also be present (Clark 2004). Evidence for this possibility comes from results that show no association between individual genome-wide measures of heterozygosity and a measure of fitness, body condition, for snakes from these populations (Chapter 3). From a conservation perspective this suggests that genetic factors may have little impact on the long-term persistence of small populations of these snakes but rather ecological factors may play a larger role in determining whether populations survive in the long term.

The unexpectedly low migration rates estimated by MIGRATE raises the question of whether these results could be an artifact of sampling or other features of the populations. Parameter estimates in MIGRATE are influenced by recent and significant declines in population size which bias long-term parameter estimates of \( N_e \) and therefore \( M \) (Beerli
2009) and the presence of unsampled “ghost” populations can inflate \( \Theta \) and thereby affect migration estimates (Beerli 2004). However, I do not believe that either factor has influenced parameter estimates in this study. First, there is no genetic evidence for recent and substantial declines in most populations (Table 2.2). Second, it seems highly unlikely that any unsampled “ghost” populations would inflate estimated migration rates based on the results from the regional analyses, demonstrating extremely limited migration between nearby populations. Finally, independent support for my general conclusion that populations are in long-term drift-migration equilibrium comes from the 2mod analysis, which strongly supported a population model consistent with a historical pattern of limited gene flow between isolated populations.

*High Levels of Population Differentiation*

Massasauga populations in this study exhibit high levels of genetic differentiation at both range-wide and regional scales which corroborates earlier findings by Gibbs *et al.* (1997) based on more limited analyses. Similar levels of differentiation between populations > 100 km apart has been documented for populations of another subspecies (*edwardsii*) of *S. catenatus* present in Arizona and New Mexico which occupy quite different habitats (short grass prairie) (Anderson *et al.* 2009) suggesting that a high level of structure may characterize this species in general. Overall, patterns of moderate to high genetic differentiation appear to be a general pattern among snakes in fragmented landscapes (Bushar *et al.* 1998; Clark *et al.* 2008; Manier & Arnold 2005; Marshall *et al.* 2009; Paquin *et al.* 2006; Prior *et al.* 1997; Prosser *et al.* 1999; Ursenbacher *et al.* 2009).
An important conservation implication suggests discrete populations of S. catenatus represent demographically independent units such that migration and recolonizations will have little effect on the dynamics of individual populations (Gibbs et al. 1997; Anderson et al. 2009). My results show that migration among populations of these snakes was likely low even before humans had significant impacts on the landscape. From a practical standpoint, this still means that each population, even those at the regional level, need to be managed as an independent conservation unit.

Variation in Effective Population Size

My estimates of genetic effective population size from the MIGRATE analyses (θ\text{Ne}) show substantial variation in the sizes of populations under study, varying from 0.16 to 2.2 for the range-wide analyses. These appear to represent long-term estimates of population sizes because with one exception (PRF), populations do not show the genetic signature of either historical or recent declines in numbers over time. These data represent the first set of range-wide estimates of genetic effective population size using a common set of genetic markers and similar method of analyses for this endangered taxon. This allows an assessment of the relative importance of different populations as potential reservoirs of standing genetic variation on a taxon-wide basis. It also allows for the identification of populations which have relatively low levels of variation and so could potentially be at increased risk of extinction due to a loss of adaptive potential (Frankham
et al. 2002). For example, the BPNP population in Ontario, and the KL and WM populations in Ohio all have the highest values for $\theta_{Ne}$ (all > 1.0) hence the highest levels of standing neutral variation among those analyzed. This feature could be used to assign additional conservation status to these populations in recognition of the fact that they possibly contain substantial amounts of the overall adaptive genetic variation for this taxon. In contrast, populations that show unusually low values for $\theta_{Ne}$ (e.g. PRF, GRL-1 and WAIN) may now be recognized as potentially being at increased risk of extinction due to loss of adaptive genetic variation.

I caution that use of $\theta_{Ne}$ to estimate levels of standing adaptive variation in eastern massasauga populations relies on a number of important assumptions. First, $\theta_{Ne}$ is assumed to be based on standing genetic variation from a single interbreeding deme of snakes. I attempted to satisfy this assumption by genotyping snakes from a single restricted location but, given the presence of fine-scale population differentiation in these animals (Gibbs et al. 1997; this study) unrecognized subpopulation differentiation may act to inflate observed levels of variation as well as values of $\theta_{Ne}$ for populations showing high values for this parameter (e.g. BPNP). Second, and most significantly, I assume that levels of neutral genetic variation in microsatellites correlate with levels of adaptive genetic variation. Whether this is the case in general is controversial (Chapman et al. 2009) but there are two pieces of evidence that this correlation may be weak in these snakes. First, population-specific values of $\theta_{Ne}$ show no significant correlation with population level variation in an adaptive trait (venom variation) that is at least partially under genetic control (Gibbs & Chiucchi, 2011; see Gibbs & Rossiter 2008; Gibbs et al.
2009 for discussion of causes of venom variation). Second, a measure of individual
fitness (body condition) is not correlated with genome-wide measures of individual
heterozygosity based on microsatellite variation (Chiucchi and Gibbs, 2011). While
limited in scope to just two types of fitness-related traits, these results caution that links
between heterozygosity and fitness may not be as strong in this taxa as those found for
other species (Reed & Frankham 2003a).

A second use for population-specific values of $\theta_{Ne}$ is to gain a rough idea of the
magnitude of the census sizes for each population. Reliable and comparable census sizes
are not available for this taxon due to the difficulties of comparing the results of different
survey methods used to estimate snake abundance in different locations (Johnson 2000).
Nonetheless, size estimates are crucial for important conservation analyses such as
Population Viability Analyses (PVA) which have been conducted for this species
(Middleton & Chu 2004; Miller 2005; Seigel & Sheil 1999). One way to qualitatively
link genetic ($N_e$) and census ($N_c$) population sizes is to use the result for a wide variety of
organisms which show that the ratio of $N_e/N_c$ is usually less than 0.5 (Frankham 1995)
with the best estimates for vertebrates averaging 0.38 (Vucetich et al. 1995). If I assume
a widely-used mutation rate of $5 \times 10^{-4}$ mutations/generation/locus (Garza & Williamson
2001) to convert estimates of $\theta_{Ne}$ to $N_e$, then estimates of $N_e$ range from ~ 80 (WAIN) to
1100 (BPNP) individuals. Assuming a ratio of $N_e/N_c = 0.38$, then I estimate that $N_c$ for
populations surveyed here range from ~ 200 to 2900 snakes. If I use a more conservative
ratio of $N_e/N_c = 0.5$ then all populations have $N_c$ values of $\geq 160$ individuals. Although
estimated values of $N_c$ are crucially dependent on the mutation rate and ratio of $N_e/N_c$
used, my results suggest that under a variety of parameter values census sizes of all populations are greater than 160 individuals. This is significant, as my estimate is greater than the minimum population size (~100 individual snakes) required to maintain a viable population of snakes over hundreds of generations according to a number of different PVA analyses conducted under a range of demographic assumptions (Middleton & Chu 2004; Seigel & Sheil 1999). I caution however, that a crucial assumption with this use of genetically-based measures of $\theta_{Ne}$ is that levels of genetic variation reflect contemporary census sizes. Of concern is that populations of eastern massasaugas in several US states are thought to have undergone recent declines over short time periods (M. Dreslik, pers. com.) and so long-term estimates of $N_e$ based on microsatellite variation may overestimate census sizes of these populations.
References


Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3) Available from [http://www2.unil.ch/popgen/softwares/fstat.htm](http://www2.unil.ch/popgen/softwares/fstat.htm).


Table 2.1. Sample size and genetic diversity estimates for each locality (ID) across 19 microsatellite loci. Observed heterozygosity ($H_O$) and expected heterozygosity ($H_E$) were calculated using FSTAT version 2.9.3. Allelic richness ($AR$) was calculated using MSA based on a minimum of six individuals. The mode and 90% highest probability density limits of $F$ as calculated with 2mod are shown.

<table>
<thead>
<tr>
<th>Regional Group</th>
<th>Population</th>
<th>ID</th>
<th>N</th>
<th>$H_O$</th>
<th>$H_E$</th>
<th>$AR$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern Illinois</td>
<td>South Shore State Park</td>
<td>SSSP</td>
<td>18</td>
<td>0.62</td>
<td>0.65</td>
<td>3.91</td>
<td>0.15 (0.12 - 0.19)</td>
</tr>
<tr>
<td></td>
<td>Eldon Hazlet State Park</td>
<td>EHSP</td>
<td>14</td>
<td>0.65</td>
<td>0.69</td>
<td>4.66</td>
<td>0.25 (0.20 - 0.29)</td>
</tr>
<tr>
<td>Western and Central Ohio</td>
<td>Prairie Road Fen</td>
<td>PRF</td>
<td>21</td>
<td>0.50</td>
<td>0.56</td>
<td>3.31</td>
<td>0.31 (0.25 - 0.36)</td>
</tr>
<tr>
<td></td>
<td>Killdeer Plains Wildlife Area 1</td>
<td>KL-1</td>
<td>68</td>
<td>0.75</td>
<td>0.77</td>
<td>5.23</td>
<td>0.07 (0.05 - 0.08)</td>
</tr>
<tr>
<td></td>
<td>Killdeer Plains Wildlife Area 2</td>
<td>KL-2</td>
<td>17</td>
<td>0.75</td>
<td>0.76</td>
<td>4.96</td>
<td>0.09 (0.06 - 0.10)</td>
</tr>
<tr>
<td></td>
<td>Willard Marsh Wildlife Area</td>
<td>WM</td>
<td>15</td>
<td>0.65</td>
<td>0.68</td>
<td>4.68</td>
<td>0.13 (0.10 - 0.16)</td>
</tr>
<tr>
<td>Northeast Ohio</td>
<td>Grand River Lowlands 1</td>
<td>GRL-1</td>
<td>18</td>
<td>0.58</td>
<td>0.57</td>
<td>3.31</td>
<td>0.35 (0.31 - 0.42)</td>
</tr>
<tr>
<td></td>
<td>Grand River Lowlands 2</td>
<td>GRL-2</td>
<td>18</td>
<td>0.54</td>
<td>0.51</td>
<td>2.90</td>
<td>0.40 (0.35 - 0.49)</td>
</tr>
<tr>
<td></td>
<td>Grand River Lowlands 3</td>
<td>GRL-3</td>
<td>20</td>
<td>0.64</td>
<td>0.63</td>
<td>3.78</td>
<td>0.25 (0.21 - 0.30)</td>
</tr>
<tr>
<td>Western Pennsylvania</td>
<td>State Game Lands 95</td>
<td>SGL95</td>
<td>6</td>
<td>0.53</td>
<td>0.58</td>
<td>3.16</td>
<td>0.35 (0.27 - 0.41)</td>
</tr>
<tr>
<td></td>
<td>Venango County</td>
<td>VEN</td>
<td>7</td>
<td>0.50</td>
<td>0.57</td>
<td>3.80</td>
<td>0.23 (0.20 - 0.31)</td>
</tr>
<tr>
<td></td>
<td>Jennings Environmental Education Center</td>
<td>JEEC</td>
<td>9</td>
<td>0.53</td>
<td>0.62</td>
<td>3.79</td>
<td>0.21 (0.17 - 0.28)</td>
</tr>
<tr>
<td>New York</td>
<td>Bergen Swamp</td>
<td>BS</td>
<td>20</td>
<td>0.50</td>
<td>0.49</td>
<td>2.86</td>
<td>0.41 (0.36 - 0.50)</td>
</tr>
<tr>
<td></td>
<td>Cicero Swamp</td>
<td>CS</td>
<td>62</td>
<td>0.51</td>
<td>0.52</td>
<td>3.08</td>
<td>0.33 (0.28 - 0.37)</td>
</tr>
<tr>
<td>Ontario</td>
<td>Ojibway Prairie</td>
<td>OJIB</td>
<td>8</td>
<td>0.48</td>
<td>0.60</td>
<td>3.91</td>
<td>0.20 (0.14 - 0.23)</td>
</tr>
<tr>
<td></td>
<td>Wainfleet Bog</td>
<td>WAIN</td>
<td>12</td>
<td>0.61</td>
<td>0.66</td>
<td>4.24</td>
<td>0.15 (0.11 - 0.19)</td>
</tr>
<tr>
<td></td>
<td>Warder Ranch, Bruce Peninsula National Park</td>
<td>BPNP</td>
<td>20</td>
<td>0.66</td>
<td>0.68</td>
<td>4.84</td>
<td>0.08 (0.06 - 0.11)</td>
</tr>
<tr>
<td></td>
<td>Beausoleil Island, Georgian Bay Islands National Park</td>
<td>GBINP</td>
<td>15</td>
<td>0.58</td>
<td>0.62</td>
<td>4.07</td>
<td>0.18 (0.14 - 0.22)</td>
</tr>
<tr>
<td></td>
<td>Killbear Provincial Park</td>
<td>KPP</td>
<td>20</td>
<td>0.61</td>
<td>0.60</td>
<td>3.71</td>
<td>0.23 (0.20 - 0.29)</td>
</tr>
</tbody>
</table>
Table 2.2. Summary of results from BOTTLENECK for 16 populations for the Wilcoxon’s test for both the two-phase model (TPM) and stepwise model (SMM) along with results from the mode-shift test. Under the mode-shift test, an L-shaped distribution of alleles is expected in the absence of a bottleneck whereas a distribution with a shifted mode is expected in a population that has gone through a bottleneck.

<table>
<thead>
<tr>
<th>Population ID</th>
<th>Wilcoxon's test TPM</th>
<th>Wilcoxon's test SMM</th>
<th>Mode-shift test</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSSP</td>
<td>0.45</td>
<td>0.74</td>
<td>L-shape</td>
</tr>
<tr>
<td>EHSP</td>
<td>0.81</td>
<td>0.67</td>
<td>L-shape</td>
</tr>
<tr>
<td>PRF</td>
<td>0.00067</td>
<td>0.00019</td>
<td>L-shape</td>
</tr>
<tr>
<td>KL</td>
<td>0.61</td>
<td>0.34</td>
<td>L-shape</td>
</tr>
<tr>
<td>WM</td>
<td>0.76</td>
<td>0.46</td>
<td>L-shape</td>
</tr>
<tr>
<td>GRL-1</td>
<td>0.18</td>
<td>0.34</td>
<td>Shifted Mode</td>
</tr>
<tr>
<td>GRL-2</td>
<td>0.71</td>
<td>0.55</td>
<td>L-shape</td>
</tr>
<tr>
<td>GRL-3</td>
<td>0.60</td>
<td>0.44</td>
<td>L-shape</td>
</tr>
<tr>
<td>JEEC</td>
<td>0.74</td>
<td>0.89</td>
<td>Shifted Mode</td>
</tr>
<tr>
<td>BS</td>
<td>0.82</td>
<td>0.70</td>
<td>L-shape</td>
</tr>
<tr>
<td>CS</td>
<td>0.34</td>
<td>0.63</td>
<td>L-shape</td>
</tr>
<tr>
<td>OJIB</td>
<td>0.71</td>
<td>0.82</td>
<td>Shifted Mode</td>
</tr>
<tr>
<td>WAIN</td>
<td>0.15</td>
<td>0.07</td>
<td>L-shape</td>
</tr>
<tr>
<td>BPNP</td>
<td>0.57</td>
<td>0.14</td>
<td>L-shape</td>
</tr>
<tr>
<td>GBINP</td>
<td>0.15</td>
<td>0.04</td>
<td>L-shape</td>
</tr>
<tr>
<td>KPP</td>
<td>0.11</td>
<td>0.04</td>
<td>L-shape</td>
</tr>
</tbody>
</table>
Figure 2.1. A) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. Population codes for the sampling locations are found in Table 2.1. Populations with asterisks are those used in the range-wide analysis using STRUCTURE. KL represents both KL-1 and KL-2. State abbreviations are as follows: IL (Illinois); IN (Indiana); KY (Kentucky); OH (Ohio); WV (West Virginia); PA (Pennsylvania); NY (New York); MI (Michigan); ON (Ontario, Canada). B) Inset showing the relative location of the sampling localities within the United States and Canada.
Figure 2.2. Bar plot for STRUCTURE at K = 11 (ln P (D) = -14834) showing the association among genetic clusters and sample locations. Abbreviations for sampling locations are as follows: SSSP (South Shore State Park); PRF (Prairie Road Fen); KL-1 (Killdeer Plains Wildlife Area); WM (Willard Marsh Wildlife Area); GRL-1 (Grand River Lowlands 1); WAIN (Wainfleet Bog); BPNP (Bruce Peninsula National Park); GBINP (Georgian Bay Islands National Park); KPP (Killbear Provincial Park); BS (Bergen Swamp); CS (Cicero Swamp).
Figure 2.3. Bar plot for STRUCTURE at K = 3 showing the association among genetic clusters and sample locations for northeastern Ohio. Abbreviations for sampling locations are as follows: GRL-1 (Grand River Lowlands 1); GRL-2 (Grand River Lowlands 2); GRL-3 (Grand River Lowlands 3).
Figure 2.4 Population estimates of theta ($\theta_N$) from MIGRATE. Values are shown in decreasing order from left to right. Error bars represent the 95% confidence intervals for each estimate. KL represents KL-1.
Figure 2.5.A. Likelihood estimates of long-term gene flow among populations at a) range-wide and b) regional scales. Circle sizes reflect relative population size. Values inside the circles are θₜ, shown with 95% confidence intervals. Bold values indicate migration was asymmetrical as determined from non-overlapping 95% confidence intervals and the arrows represent the direction of gene flow among sites. All pairwise estimates of M (m/µ) are shown with 95% confidence intervals. Black bars represent the position of the Great Lakes in relation to each population.
Figure 2.5.B. Likelihood estimates of long-term gene flow among populations at a) range-wide and b) regional scales. Circle sizes reflect relative population size. Values inside the circles are $\theta_0$, shown with 95% confidence intervals. Bold values indicate migration was asymmetrical as determined from non-overlapping 95% confidence intervals and the arrows represent the direction of gene flow among sites. All pairwise estimates of $M$ ($m/\mu$) are shown with 95% confidence intervals. Black bars represent the position of the Great Lakes in relation to each population.
Figure 2.6. Distribution of differences between contemporary migration rates ($m_c$) estimated using BAYESASS and historical migration rates ($m_h$) estimated using MIGRATE for specific population comparisons ($m_c - m_h$) as shown in Figure 2.4.
Chapter 3: Lack of a heterozygosity-fitness correlation in isolated populations of a threatened rattlesnake, the eastern massasauga (*Sistrurus c. catenatus*)

Abstract

Endangered species often occur in small isolated populations with greater risk of short-term extinction due to factors such as the negative effects of inbreeding depression. The eastern massasauga (*Sistrurus c. catenatus*) is an endangered snake that is found in isolated populations of varying size with little to no opportunities for gene flow. As such these populations may have a history of inbreeding and possibly suffer from the negative effects of inbreeding. Here, I investigate this possibility by 1) Testing for a heterozygosity-fitness correlation using a measure of individual multilocus heterozygosity (MLH) based on 19 microsatellite loci and an indirect measure of fitness (body condition) from snakes in 14 populations; 2) Assessing whether there is any evidence of inbreeding by comparing observed mean MLH for each population with the expected distribution under random mating using a permutation procedure and; 3) Comparing the estimates of body condition in wild snakes to that of snakes housed in zoos which I use as a benchmark for individuals in good condition. I found significant variation in both mean MLH and mean body condition among populations but no correlation between MLH and body condition even after controlling for non-genetic factors such as sex, season of capture and year of capture which suggests a lack of
inbreeding depression. The permutation of MLH values suggests random mating in all but one population implying a lack of inbreeding in these snakes. Wild snake body condition is on average comparable with captive snakes with some wild populations in better condition and some in worse condition. Overall, my results suggest among-population differences in body condition cannot be explained by genetics but that ecological factors may have a greater effect on individual fitness in these snakes at least as measured as body condition.

Introduction

The relationship between fitness and heterozygosity is an important one to explore because of the potential impact on extinction of endangered species (Frankham 2005; Oostermeijer et al. 2003; Reed & Frankham 2003b) particularly those with isolated populations due to severely fragmented habitat (Lopez-Pujol et al. 2005). This can occur through the exposure of recessive deleterious alleles through inbreeding (Lacy 1992; Ledig 1986; Ralls & Ballou 1983; Storfer 1999). Often, species that fit this criterion are endangered and occur in small populations that can increase the chance for inbreeding and can lead a population to experience the effects inbreeding depression. This effect has been observed in many taxa including two species of snakes: the meadow viper in Hungary (Ujvari et al. 2002) and the adder in Sweden (Madsen et al. 1996). In meadow vipers, inbreeding depression is evident in reports of birth deformities, chromosomal abnormalities, low juvenile survival and reduced variation in major histocompatibility (Mhc) class I loci (Ujvari et al. 2002). Highly isolated populations of adders in Sweden
exhibit the effects of inbreeding depression through smaller litter size relative to maternal body size, deformed and/or stillborn offspring, and lower levels of genetic diversity among individuals (Madsen et al. 1996). However, not all animals found in small isolated populations show the effects of inbreeding depression. For example, the endangered copper redhorse (*Moxostoma hubbsi*) occurs in a short section of the St Lawrence River in Canada and populations have been fragmented and declining since 1849. Relatedness values suggest an outbred population rather than inbred populations that the authors attribute to long generation times for this species (Lippe et al. 2006). Similar results were found in wood ants where fragmented populations show high levels of genetic variation and no signs of inbreeding likely due to their polygynous mating system (Gyllenstrand & Seppa 2003).

Species in fragmented habitats occurring in isolated populations often suffer from low migration rates between populations thereby potentially increasing the effects of inbreeding. An example of a species where this is potentially an issue is the eastern massasauga rattlesnake (*Sistrurus c. catenatus*) that occurs in a series of small isolated populations with little to no opportunities for gene flow and has persisted this way for the last few thousand years (Chiucchi & Gibbs 2010). Given the long-term isolation of these populations, it seems possible that these animals could suffer from the potential fitness costs associated with prolonged inbreeding. However, it is also possible that these animals do not suffer from the costs of inbreeding possibly because they avoid inbreeding through a mechanism such as kin recognition (Clark 2004) or because repeated exposure to non-random mating has led to the exposure of deleterious alleles to selection thereby
eliminating the deleterious effects of inbreeding from the population (Keller & Waller 2002).

Evidence of inbreeding and the effects of inbreeding depression on fitness can be studied by examining the relationship between individual multilocus heterozygosity (MLH) and indirect measures of fitness (e.g. body condition) and are often referred to as genetic diversity fitness correlations (GDFC’s or GFC’s) or heterozygosity fitness correlations (HFC) (Chapman et al. 2009). Many studies have shown a significant positive relationship between fitness and MLH from microsatellite loci (Amos et al. 2001; Bean et al. 2004; Coltman et al. 1998; David 1998; Gage et al. 2006; Hansson et al. 2001; Slate et al. 2000) though the strength of GFC’s in natural populations is still highly debated (Da Silva et al. 2006; Grueber et al. 2011). Often, stronger correlations are observed when using life history traits such as age at first reproduction or average litter size rather than morphological traits such as body condition (Coltman & Slate 2003; DeRose & Roff 1999) though significant correlations between morphological traits and heterozygosity have been observed (Curik et al. 2003).

Estimating fitness in wild populations has been done using a variety of indirect measures (Jakob et al. 1996). In reptiles, one indirect measure that has been commonly used is to take an animal’s mass and length and estimate body condition as a scaled measure of individual mass relative to length (Jakob et al. 1996). The most common way to estimate body condition in reptiles is to use the residuals from a regression of mass on snout-vent length, known as the Residual Index. This has been calculated for collard lizards (Husak
2006), tuatara (Moore et al. 2007), and various snakes (Moore et al. 2000; Reading 2004; Sperry & Weatherhead 2008; Weatherhead & Brown 1996). Recently though, this approach has been challenged, with critics pointing out the inability of the residual index to accommodate the scaling relationship between mass and length (Peig & Green 2010). The Scaled mass index accounts for this scaling issue and has recently been shown to be a more reliable estimate of body condition when compared to the residual index and is particularly useful because it allows for direct comparison across populations because it retains the original Y-axis units (Peig & Green 2009).

Here, my primary goal was to explore whether there are genetic fitness costs to massasaugas living in small, isolated populations for thousands of years. I do this in two main ways. First, I test for evidence of inbreeding depression by investigating the relationship between fitness (estimated as relative body condition) and genetics (multilocus heterozygosity) among 14 populations of the eastern massasauga rattlesnake. Next, I test for inbreeding and increased levels of relatedness using a permutation procedure. Finally, I compared estimates of body condition from wild populations with values from reference individuals in the form of captive animals housed in zoos.

Methods

Data collection and Estimates of Body Condition

With the help of many colleagues, a total of 248 snakes from 14 populations throughout the range of the massasauga were sampled (Figure 3.1, Table A.3). For each snake, mass
(in grams) and snout-vent length (SVL, in mm) was recorded along with sex, female gravid status, season of capture and date of capture. Body condition was used as an estimate of fitness with more “fit” individuals considered to be those in better condition. To do so, I calculated the Scaled Mass Index (Peig & Green 2009) that seeks to standardize mass to a specific fixed body length and is designed to incorporate allometric changes in scaling that might be observed in many species. This index has recently been shown to outperform older methods of determining body condition from mass and length estimates because of its use of model II linear regression, specifically standardized major axis regression, which incorporates the likelihood that both variables have some underlying error rate associated with their measurement (Peig & Green 2010). A major advantage of this technique is that it creates standardized Y-axis values that retain their original units and scale and are therefore easily compared across populations (Peig & Green 2010). To calculate the scaled mass index, individuals across all populations were included in the original regression and specific details about the calculation can be found in Peig and Green (2009). Briefly, both mass and SVL were log-transformed and Model II Regression in R version 2.11.1 (R Development Core Team 2010) was used to calculate the slope ($b_{SMA}$) of the best-fit line from a standardized major axis regression. $L_o$ was calculated as the mean SVL from the entire dataset (mean SVL = 497 mm). I combined all treatment groups into one analysis under the assumption that snakes in different treatment groups store fat and energy reserves in similar ways and so allometric relationships between length and weight are similar. A small number of gravid females were removed from all analyses because they are expected to be heavy animals relative to their length and can potentially bias the relationship of mass to SVL for the species.
I was concerned about estimating the impact of variation among individual researchers and in terms of methods of measurement on estimates of body condition. Researchers collecting data used two different techniques to estimate SVL for snakes in this study: 1) the tube and 2) squeeze box methods. The “tube” method involves placing the anterior portion of the snake into a see-through tube and using seamstress tape to follow the length of the animal’s body to obtain the SVL. The “squeeze box” method entails use of a squeeze box (Bertram & Larsen 2004) which traps the animal under a see-through piece of Plexiglas and allows the researcher to trace the outline of the animal with a marker. Next, the traced line is measured using seamstress tape. I estimated both among-researcher and among-technique error in measurement of body condition in the following way. Two researchers, who had contributed data to this project, recorded measurements from a captive population of 15 individuals from a congeneric species (pygmy rattlesnakes [Sistrurus miliarius]) using the tube method and the squeezebox method. The mass and SVL of each person’s measurements were then converted into a scaled mass index as described above. For this analysis, I calculated a new $b_{SMA}$ because this is a different species and so might have a different slope to describe their growth pattern. Next, individual Scaled Mass Index estimates were grouped by researcher and technique and paired t-tests were used to determine differences between researcher and technique. I then compared the magnitude of difference between researchers and technique to the observed differences in Scaled Mass Index among my populations. This allowed us to determine if error among researchers and technique could account for the differences I observed among populations (see below).
Estimating genetic diversity

Samples were genotyped at 19 microsatellite loci as detailed in Anderson et al. (2010) and loci were checked for null alleles as described in Chiucchi and Gibbs (2010). Individual multilocus heterozygosity (MLH) was measured as the total number of heterozygous loci divided by the total number of genotyped loci for a given individual (Slate 2004). Most individuals were genotyped at 19 loci however, a minor number of individuals (n=17 [7%]) were genotyped at 18 of the 19 loci. For population comparisons, mean MLH was calculated as the average of the individual heterozygosity estimates for a given population.

Genetic Fitness Correlations

General linear mixed models (GLMM's) were used to test which variables explained variation in body condition the best (Zuur et al. 2009). In all models, body condition was the response variable with heterozygosity (as individual MLH), season of capture, year, and sex treated as independent variables. Season of capture, year and sex were all included as independent variables because they all could potentially be responsible, at least in part, for any differences in body condition. For example, snakes captured in spring could weigh less than those captured in summer simply because they have just emerged from hibernation and have not had the opportunity to feed yet where as snakes captured later in the year (summer or late summer) would likely have had opportunities to feed and gain weight. For this reason, “season of capture” was broken up into three categories, spring (April and May captures), summer (June and July captures) and
autumn (August and September captures). The year in which snakes were captured can be influenced by many environmental factors, such as drought and high summer temperatures, that may have varied on a yearly basis and can affect snake body condition so I treated “year of capture” as an independent variable in this analysis. Massasaugas display sexual size dimorphism with males on average obtaining larger size than females so I included “sex” as an independent variable as well. Initially, all independent variables were included in the model (full model) along with their interaction terms. However, all interaction terms were non-significant so were removed them from further analysis. Independent variables were removed sequentially as determined by Akaike’s Information Criterion (AIC) until removal of a given term failed to improve model fit (minimal model) (Zuur et al. 2009). In all models, population was treated as a random effect. Models were estimated using restricted maximum likelihood (REML) methods and $F$ and $P$ values from the full and minimal models are reported. Likelihood ratio tests were used to determine the significance of random effects by comparing models with and without random effects included. All tests were performed using the nlme package in R version 2.11.1 following guidelines in Zuur et al. (2009).

**Estimating Inbreeding**

As inbreeding increases within a population, individuals are expected to have increased levels of genome-wide homozygosity and therefore decreased MLH estimates compared to a randomly breeding population (Slate et al. 2004). To estimate levels of inbreeding I calculated the MLH for each individual as described above and then calculated the mean
MLH value for each population. I then compared the observed mean MLH to the expected distribution assuming a random mating population using a permutation test as described by Castric et al (2002). To do so, I used MATLAB (2010R) to simulate 1000 randomly mating populations by permuting alleles within a locus and then calculating the mean MLH for each simulated population. If a population is experiencing significant inbreeding, the observed mean MLH of my populations is expected to be less than 95% of the permuted mean MLH values. A p-value was then calculated as the number of simulated values less than the observed mean MLH values and significance was determined at the 95% level after correcting the level of significance for multiple tests using the B-Y method (Benjamini & Yekutieli 2001; Narum 2006).

**Body Condition Estimates from Zoo Populations**

In an attempt to establish a benchmark body condition, I obtained estimates of mass and SVL from 22 massasaugas housed at six zoos across the country (Table A.4). I assumed that these represent well-fed healthy individuals with “good” body condition (but see below). I then added these animals to the linear regression analysis of mass and SVL in order to estimate the body condition for the “zoo population.” Finally, I compared wild snake scaled mass index estimates with those of zoo snakes by using the estimated value of $b_{SMA}$ and mean SVL from the initial model II regression of the total population above. This allowed us to directly compare the zoo snakes to those in the wild populations.
Results

A total of 262 snakes were sampled from 14 wild populations with sample sizes varying from 10 to 50 individuals per population. The relationship between log-transformed mass and log-transformed SVL was highly linear ($r^2 = 0.932, p = 0.000$, Figure 3.2).

The full GLMM resulted in a lower AIC value than any reduced models so I only report the results from this model. Overall, GLMM’s failed to detect any evidence of inbreeding depression: MLH and all other independent variables (sex, age and time) were not significant predictors of body condition in all models tested (Table 3.1). However, there was a significant effect of population (Table 3.1, $p = 0.0001$) with snakes in certain populations have higher relative body condition (PRF, GRL-2, CS and BPNP) while others had lower condition (EHSP, SSSP and KPP). The average population scaled mass index ranged from 142.3 grams to 211.4 grams and ninety-five percent confidence intervals for mean population body condition revealed seven populations to be in average condition, four populations above average and three populations below average (Figure 3.3). ANOVA’s revealed significant differences among populations for average body condition ($F = 4.89, p < 0.001$, Figure 3.3) as well as individual heterozygosity ($F = 12.75, p < 0.000$; Figure 3.4).

Paired t-tests for researcher error were not significant ($p = 0.28$) with the mean difference in condition measures between researchers of 1.41 and 1.37 grams per cm (gpc) for the tube method and the squeezebox method, respectively. Tests for difference among technique were significantly different ($p < 0.001$); however, despite this significant
difference, the mean difference in Scaled Mass Index using different techniques was very small (1.85 and 1.87 gpc for researchers 1 and 2, respectively). This is substantially lower than the observed mean Scaled Mass Index differences among populations in this study (range: 142.3 – 211.4 gpc; SD: 38.7 grams) and suggests that these differences are not simply a result of measurement error.

Mean MLH ranged from 0.497 in PRF to 0.761 at KL. After correcting for multiple tests, evidence for inbreeding as determined by a shift towards lower MLH values than expected, was observed only in PRF (mean MLH = 0.529; p = 0.01). All other populations had mean MLH estimates within the range of values expected for a randomly mating population (Table 3.2).

Finally, snakes from zoos exhibited average body condition (mean = 173.2 gpc) varying only slightly from the mean scaled mass index for wild snakes (all snakes pooled, mean = 175.72). In addition, the variance associated with the zoo snakes (SD = 37.24) was similar in magnitude to those of the wild snakes (SD = 39.07) (Figure 3.3).

Discussion

My major results are (i) there is no evidence of inbreeding depression as indicated by a lack of a correlation between MLH and body condition in these snakes; (ii) body condition is variable among populations and this variation cannot be explained by genetics, sex, season of capture or year of capture; and (iii) snakes within each population
appear to be mating at random with no evidence of inbreeding or avoidance of inbreeding. Below, I discuss the evolutionary and ecological implications of these results.

No evidence of a genetic-fitness correlation

My investigation into the effects of MLH on body condition in these populations failed to detect any significant relationship even after potential non-genetic effects (sex, year, season of capture) were taken into account. Overall, these results indicate snakes in good condition show no difference in levels of MLH compared to snakes in relatively poor condition suggesting there is no advantage to an increased level of heterozygosity as measured using microsatellite loci in eastern massasaugas.

The lack of any evidence of inbreeding depression in these populations is somewhat surprising given the small sizes of some of these populations and length of time that they have been isolated (Chiucchi & Gibbs 2010). This suggests one of two possible explanations; a lack of inbreeding possibly through behavioral avoidance mechanisms such as kin recognition (Clark 2004) or repeated exposure of deleterious alleles from mating between closely related individuals and subsequent purging of genetic load via genetic drift (Glemin 2003). From my results, this second possibility is unlikely, as it should also have produced a signal of inbreeding in these populations. Most populations had MLH values within the expectations of a randomly mating population that suggests inbreeding does not occur. This has important implications because it suggests that
although these populations vary by effective population size, populations remain large enough to prevent any serious levels of inbreeding. The genetically-effective population size of these snakes ranges from a minimum of 80 (WAIN) to a maximum of 1100 individuals (BPNP) with estimated census sizes of 160 to >2000 individuals (Chiucchi & Gibbs 2010). Assuming these estimates are correct, these animals are able to avoid inbreeding and the effects of inbreeding depression despite population sizes that are less than 200 individuals in some cases.

The best-known examples of inbreeding in snakes come from adder populations in northern Europe (Madsen et al. 1996) and Antiguan racer from the small island of Antigua (Daltry et al. 2001). Both species have been shown to suffer some extreme effects of inbreeding depression but each has a population size that is smaller than almost all massasauga populations in this study. The European adders have an estimated population size (census size) of approximately 40 animals (Madsen et al. 1996) whereas the Antiguan racer population was last estimated to harbor only 80 animals (Daltry et al. 2001). In both of these populations, the loss of genetic diversity due to inbreeding, likely in combination with genetic drift, has led to severe morphological deformities. Further, when inbreeding depression is observed it is often done so through a reduction in growth rate, fertility, fecundity and resistance to disease (Charlesworth & Charlesworth 1987). Currently, there is no evidence for these issues in massasauga populations. My smallest population (WAIN, estimated census size = 160 animals) is approximately 4 times the size of the European adder population studied by Madsen et al. (1996) and shows no evidence of inbreeding depression. Because massasaugas have been isolated for
thousands of years without any detectable negative effects of inbreeding, maintaining current effective population sizes of massasaugas will likely be an effective strategy in minimizing inbreeding depression in these populations.

*Geographic variation in body condition*

Another important and very interesting result from this study is that populations show high levels of variation in body condition as determined by the scaled mass index that directly compares mass per unit length. I see three potential mechanisms that could lead to this result. First, measurement error could potentially create the observed pattern in body condition but the results from my comparison of researcher and technique effect suggest it is unlikely this variation can create differences in body condition as large as those observed in my study populations. The difference in mean body condition among researchers was 1.4 gpc accounting for approximately 2% of the largest observed difference between my populations. Furthermore, the difference in mean body condition among techniques was 1.8 gpc, which accounts for only 2.6% of the largest difference among my populations. From these results, I conclude that differences in body condition among populations of massasauga are real; researcher and technique error cannot create differences in body condition of the magnitude that I observe in these populations.

Before discussing the biological reasons for variation in body condition it is valid to ask whether or not my estimates of body condition actually measure body condition and amount of available fat stores in snakes. Previous work has documented that
approximately 50% of the variance in body condition estimates is due to variation in stored body fat in snakes (Weatherhead & Brown 1996) while the other 50% (unexplained variance) is likely stored in places other than fat bodies, for example in liver and/or muscle tissue (Madsen & Shine 2002). Assuming massasaugas allocate energy stores to fat bodies in similar ways as other snake species, my estimates of body condition should accurately reflect the amount of fat an individual snake contains and should be nearly as reliable an estimate of body condition as actual measurements of an individual's percent fat (Weatherhead & Brown 1996). Having presented this evidence for the accuracy of my body condition estimates, I further discuss the biological implications of geographic variation in body condition and possible mechanisms that may generate this pattern in wild snake populations.

**Temporal Variation and Fluctuations in Body Condition**

Among the biological components potentially responsible for variation in body condition, I recognize two possibilities that are not necessarily mutually exclusive, temporal variation and spatial variation in ecological conditions. Temporal variation, via seasonal or annual fluctuations in body condition, has the potential to create the observed pattern in body condition seen in this study. Body condition has been shown to vary seasonally and annually in other reptile species (Moore et al. 2007) as well as in snakes (Moore et al. 2000; Sperry & Weatherhead 2008). For example, massasauga populations may experience natural fluctuations in body condition due to fluctuations in particular environmental parameters (snake density and/or prey abundance) and my sampling
scheme simply shed light on this pattern because I sampled populations at different points in their cycle. Evidence against this hypothesis comes from results from my linear mixed effects model that suggest body condition doesn’t vary annually or seasonally within this species. Data was collected throughout the active season (April-October) over many years (1994-2007) from various populations and so I was concerned my sampling methods could potentially be responsible for the observed differences in body condition among populations however I see no evidence for any variation across years or seasons.

*Spatial Variation in Ecological Parameters – Differences in Available Resources*

After ruling out methodological errors, genetic effects, and temporal variation, an alternative hypothesis is that differences in the environment (ecological effects) are likely responsible for the observed differences in body condition. For example, one potentially important environmental factor not included in the analysis capable of effecting body condition is variation in snake diet possibly stemming from variation in abundance and diversity of small mammals at each population. Evidence for this comes from results that show among-population level variation in diet using stable isotope techniques (Chapter 4) in addition to studies of other snake species showing a correlation between body condition and small mammal abundance (primary prey species) (Reading 2004; Sperry & Weatherhead 2008). Shrews represent an important group of prey for eastern massasaugas (Chapter 4) and data from one species, the masked shrew (*Sorex cinereus*), suggests their abundance varies across habitats within the state of Ohio (Pietkiewicz and Harder, in prep). If this is true, this could have a significant impact on the condition of
snakes in each population by providing the opportunity to feed more often in locations with higher shrew (or small mammal) densities. An increase in feeding opportunities would provide snakes with the ability to allocate resources into energy reserves thereby increasing fitness in these animals.

*Spatial Variation in Ecological Parameters – Differences in Density-Dependent Factors*

Density-dependent factors for example, prevalence of disease infection, may also be responsible for differences in body condition among populations. Recently, massasaugas in southern Illinois have been shown to contain high levels of infection with ophidian paramyxovirus (Allender *et al.* 2008; Allender *et al.* 2006). Two of the populations in this study with below average body condition are located in southern Illinois, EHSP and SSSP. Whether the lower observed body condition in these populations is due to the presence of ophidian paramyxovirus is unknown although this virus is known to have severe lethal issues in captive populations (Jacobson 1993). Though I do not have data on infection rates for other populations, it is reasonable to believe that rates of infection may be highly variable among populations with high rates of infection occurring at certain sites and low infection rates in others. If so, a reasonable hypothesis could postulate that variation in infection rates are responsible for the observed differences in body condition among populations.
Implications of Variation in Body Condition

Body condition has been shown to vary in other snake species and evidence suggests it may play an important role in determining if female snakes will reproduce successfully. There is empirical evidence suggesting a minimum threshold for body condition exists whereby females below this threshold fail to become reproductively active and females above the threshold tend to successfully reproduce (Naulleau & Bonnet 1996). This can have even greater effects in populations that are extremely isolated like the eastern massasauga because many population sizes are relatively small with little to no chance of recruitment via migration from other populations (Chiucchi and Gibbs 2010). If poor body condition results in females losing opportunities to breed this can have severe impacts to the overall health of the population (Gregory 2006). Female massasaugas do not breed every year but rather every second or third year, which results in a mating system where few females breed each year (Reinert 1981) hence a population with many females below the body condition threshold may decrease reproductive effort leading to a reduction in population birth rate and ultimately lowering the overall fitness and health of the population. More work is needed to determine if body condition is an important determinant for female reproduction in massasaugas and future work focusing on this question would be valuable to the long-term conservation of this species.

Limitations of the study

Though there was no relationship observed between individual genome-wide heterozygosity and a fitness related trait, body condition, at both the individual level and
the population level, the possibility of a relationship cannot be ruled out for three main reasons. First, many morphological traits are expected to be under strong stabilizing selection around some optimum rather than strong directional selection like life history traits such as clutch size (Price et al. 1988) and age at first reproduction (Jones et al. 2008; Olsen et al. 2004). Though it is possible that measuring variation around a life-history trait rather than a morphological trait may reveal a relationship between heterozygosity and fitness in this species, it is unlikely. Chapman et al. (2009) did find higher mean effect sizes in GFC studies that used life history traits rather than morphological traits, although the mean effect sizes were of similar positive magnitude. Second, if inbreeding occurs in any populations, individuals with higher inbreeding coefficients might be selected against at a very early age and simply not sampled in this study because selection has already acted. This would create a situation in which the relationship appears to be non-existent when in fact heterozygosity was measured after selection had already operated. Therefore, condition as it applies to this study is a relative term. It is entirely possible that surviving snakes measured in this study are all in “good” condition and I am simply seeing the variation around that condition. However, there were individual snakes with low heterozygosity measures (4/19 loci (21%) heterozygous) actually having one of the highest estimated values for body condition (279.4 grams). Measuring heterozygosity values this low in the dataset argues against the idea of strong selection against homozygous individuals and suggests that perhaps individuals with high inbreeding coefficients aren’t being selected against and were in fact included in this study.
Finally, it has been suggested that heterozygosity estimated from a modest number of microsatellite loci may not be a reliable estimator of genome-wide heterozygosity and therefore may be a poor predictor of inbreeding (Balloux 2004; DeWoody & DeWoody 2005; Slate et al. 2004) and that upwards of 200 microsatellite loci are needed to detect any evidence of inbreeding depression using these methods (Grueber et al. 2011). Despite the above pitfalls, the microsatellite loci used in this study were developed specifically for eastern massasauga rattlesnakes (Anderson et al. 2010), exhibit low levels of null alleles (less < 5% for most loci) and adhere to the stepwise mutation model assumed for these loci (see Chiucchi and Gibbs 2010 for details). Chapman et al. (2009) conducted a meta-analysis of GFC studies and recommended using MLH as the measure of genetic variability because it is the simplest measurement with the least number of assumptions. Here, I used the most effective estimate of genetic variability yet still find no evidence for a GFC in massasaugas though I do detect significant differences in body condition among populations.

*Body Condition of Captive Animals in Zoos*

The average body condition of the snakes held in zoos (173.2) is nearly identical to the overall mean body condition (175.7 grams) when all individuals are considered. If zoo snakes are considered as a baseline for “good” body condition, then this result implies many wild snake populations are in overall good condition while others are below the baseline, in relatively poor condition. The above assumption seems a valid one because the captive animals used in this study are currently being cared for using a strategy
designed to maintain current adult body weight (Jennifer D’Agostino, DVM, Director of Veterinary Services, Oklahoma City Zoo, pers. comm.). This result is also surprising from the perspective of captive breeding programs in zoos as well because many keepers are concerned that captive animals were being fed too often and the animals were potentially overweight (John Adamski, pers. comm.). Rather, my results suggest captive snakes are being cared for adequately, and perhaps more importantly, in a way that is typical of the prey intake of a wild snake.

Conclusions

Though my results appear to suggest that genetic factors are less important than ecological factors, this is not to say that genetics should be ignored entirely. As mentioned above, many populations have small effective population sizes that may be just large enough to prevent any significant level of inbreeding and any significant decreases in population size may be enough to drastically increase the prevalence of inbreeding thereby leading to increased levels of inbreeding depression. This being said, the conclusion that ecological factors are of great importance for this species is quite encouraging because management decisions can have substantial impacts on various ecological factors, for example, by managing for increased small mammal abundance and diversity or attempting to diagnose and treat the effects of ophidian paramyxovirus if found within the population of interest. I suggest future work focus on determining the ecological factors that may be most important for these snakes in terms of body condition.
and management decisions should be designed in a way that most effectively targets those factors.
References


Table 3.1: Results from the general linear mixed effects model with body condition as the response variable and heterozygosity, sex, season of capture and year of capture as the independent variables. Population of capture was included as a random effect and its significance was determined using a likelihood ratio test by comparing models with and without the random effect included.

<table>
<thead>
<tr>
<th></th>
<th>Full Model (AIC = 2508.54)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterozygosity (MLH)</td>
<td>$F_{1,224} = 0.215$</td>
<td>0.643</td>
</tr>
<tr>
<td>Sex</td>
<td>$F_{1,224} = 0.069$</td>
<td>0.794</td>
</tr>
<tr>
<td>Season of Capture</td>
<td>$F_{2,224} = 2.061$</td>
<td>0.129</td>
</tr>
<tr>
<td>Year of Capture</td>
<td>$F_{6,224} = 1.981$</td>
<td>0.083</td>
</tr>
<tr>
<td>Population (Random Effect)</td>
<td>$L = 15.97$</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Table 3.2: Summary of tests for inbreeding (MLH) calculated as the proportion of heterozygous loci per individual and averaged for individuals within a given population. 95% CI's were constructed by randomizing alleles for each locus to create 1,000 permuted populations from which mean MLH was calculated for each.  P-values were calculated as the number of permutations less than the observed mean value for each population. Significant values after correction for multiple tests are in bold.

<table>
<thead>
<tr>
<th>Population</th>
<th>Observed Mean</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHSP</td>
<td>0.687</td>
<td>(0.651-0.746)</td>
<td>0.32</td>
</tr>
<tr>
<td>SSSP</td>
<td>0.654</td>
<td>(0.620-0.710)</td>
<td>0.30</td>
</tr>
<tr>
<td>PRF</td>
<td>0.529</td>
<td>(0.532-0.619)</td>
<td><strong>0.01</strong></td>
</tr>
<tr>
<td>KL</td>
<td>0.767</td>
<td>(0.763-0.804)</td>
<td>0.06</td>
</tr>
<tr>
<td>WM</td>
<td>0.689</td>
<td>(0.648-0.737)</td>
<td>0.42</td>
</tr>
<tr>
<td>GRL-1</td>
<td>0.596</td>
<td>(0.514-0.611)</td>
<td>0.93</td>
</tr>
<tr>
<td>GRL-2</td>
<td>0.543</td>
<td>(0.454-0.543)</td>
<td>0.97</td>
</tr>
<tr>
<td>GRL-3</td>
<td>0.647</td>
<td>(0.583-0.667)</td>
<td>0.81</td>
</tr>
<tr>
<td>JEEC</td>
<td>0.562</td>
<td>(0.549-0.679)</td>
<td>0.07</td>
</tr>
<tr>
<td>BS</td>
<td>0.560</td>
<td>(0.461-0.547)</td>
<td>0.99</td>
</tr>
<tr>
<td>CS</td>
<td>0.542</td>
<td>(0.521-0.571)</td>
<td>0.34</td>
</tr>
<tr>
<td>WAIN</td>
<td>0.634</td>
<td>(0.611-0.722)</td>
<td>0.11</td>
</tr>
<tr>
<td>BPNP</td>
<td>0.700</td>
<td>(0.661-0.744)</td>
<td>0.50</td>
</tr>
<tr>
<td>KPP</td>
<td>0.628</td>
<td>(0.564-0.650)</td>
<td>0.83</td>
</tr>
</tbody>
</table>
Figure 3.1. A) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. Population codes for the sampling locations are as follows: EHSP (Eldon Hazlet State Park); SSSP (South Shore State Park); PRF (Prairie Road Fen); KL (Killdeer Plains Wildlife Area); WM (Willard Marsh Wildlife Area); GRL-1 (Grand River Lowlands 1); GRL-2 (Grand River Lowlands 2); GRL-3 (Grand River Lowlands 3); PA (Western Pennsylvania); BS (Bergen Swamp); CS (Cicero Swamp); WAIN (Wainfleet Bog); BPNP (Bruce Peninsula National Park); KPP (Killbear Provincial Park). B) Inset shown in the relative position of the sampling locations within the United States and Canada.
Figure 3.2. Linear relationship between log-transformed mass and log-transformed SVL for all individuals from wild populations as determined by a simple linear regression. These values were then used in a type II regression analysis (standard major axis regression) to calculate scaled mass index as described above. The dashed line represents the mean value (175.5 grams).
Figure 3.3. Mean body condition among wild populations and zoo samples of *Sistrurus catenatus*. The dashed line shows the mean scaled mass index of 175.17 grams. When using this value as the baseline for body condition, two populations, BPNP and GRL-2, are in above average condition and three populations are in below average condition (EHSP, SSSP and KPP) based on non-overlapping 95% CI (black bars) with the mean.
Figure 3.4. Mean multilocus heterozygosity (MLH) among all populations of *Sistrurus catenatus*. Bars represent 95% confidence intervals. Populations are arranged by geographic region within the range of this species from Southwest to Northeast for easier comparison of geographically close populations.
Chapter 4: Individual and Population Variation in Diet of Massasauga Rattlesnakes 
\textit{(Sistrurus catenatus)} Inferred from Stable Isotope Analysis

Abstract

Foraging behavior and diet are fundamental aspects of a species’ ecology yet are often difficult to determine for secretive animals such as snakes. Traditionally, snake diets have been estimated using gut content and fecal sample analysis but these methods have a number of potential biases in terms of determining what prey are consumed. Recent developments in stable isotope analysis (SIA) have provided researchers with new and powerful tools to conduct dietary studies that overcome some of these biases. Here, I use SIA analysis of carbon ($^{13}\text{C}$) and nitrogen ($^{15}\text{N}$) isotopes to investigate intraspecific variation in diet over time and space for an endangered snake, the eastern massasauga rattlesnake \textit{(Sistrurus c. catenatus)} for which individual- and population level-variation in diet-related traits (body condition and venom composition) has been shown. First, I used a novel Bayesian mixing model to test among eight hierarchically structured models to assess the relative importance of among-population or -individual variation in diet in these snakes. Then I employed a fully Bayesian mixing model to ascertain the contribution of individual prey items to the diet on a population-specific basis. I also calculated Manly’s alpha selectivity index that compares diet presence with availability.
for mammalian species from each of three populations of eastern massasauga. Finally, I report results using the fully Bayesian model of diet estimation for a geographically disjunct population of western massasauga rattlesnakes (*Sistrurus c. tergeminus*) in Colorado and compare my results with those for the eastern massasauga. A model identifying among-population variation but no among-individual variation had the best fit to the isotope data suggesting the importance of population level differences in diet within this species. For the eastern massasauga, shrews (*Blarina brevicauda* and *Sorex cinereus*) and voles (*Microtus pennsylvanicus*) were identified as comprising a significant component (range: 19-56%) of diet in three of five populations while mice (*Peromyscus sp.*) were rarely consumed (< 10%). Manly’s alpha selectivity index revealed a preference for shrews in two of three populations, a preference for mice in the third and an avoidance of voles in all three populations. Finally, SIA confirms previous gut content based studies showing that the diet of Western massasaugas is distinct from that of the eastern subspecies, being composed primarily of insects (33%), mammals (22%) and lizards (17%). These results suggest diet is variable across the range of these snakes and that the importance of prey items varies by population with shrews often representing an important and preferred component of the overall diet. Stable isotope analysis appears to be a viable method for estimating diet in a cryptic species such as this and these analyses represent an important step in the development of methods that can be used to determine if dietary differences are responsible for population level differences in body condition and/or venom composition.
Introduction

Diet is a fundamental aspect of the ecology and natural history of a species (Greene 1983). Intraspecific variation in diet has been observed in a number of vertebrates including fish (Binning & Chapman 2010), mammals (Stamp & Ohmart 1978), frogs (Araujo et al. 2009) and lizards (Perry 1996). Among reptiles, the patterns of intraspecific variation have been particularly well studied in snakes where diet has been shown to vary geographically (Creer et al. 2002; Daltry et al. 1998; Kephart 1982; King 1993), seasonally (Garcia & Drummond 1988; Kephart & Arnold 1982; Willson et al. 2010), ontogenetically (de Queiroz & Henke 2001; Luiselli & Agrimi 1991) and among males and females (Bowen 2004; Houston & Shine 1993). Understanding the patterns of variation in diet has provided the opportunity to gain insights into key aspects of snake biology including foraging ecology (Godley 1980), predator-prey co-evolution (Williams et al. 2004), venom evolution (Daltry et al. 1996) and the evolution of specialized morphological structures (Greene 1983).

Although the above patterns of variation in snake diet appear reasonably well documented (Luiselli 2006) there are drawbacks related to the methods used to determine diet as well as significant gaps in our knowledge of basic dietary ecology. More specifically, gut content (Shepard et al. 2004) or fecal sample analysis (Weatherhead et al. 2009) is often used to estimate diet. However these methods are limited in that they only reveal an individual’s diet over a recent short-term period (usually their most recent meal) and thus may not be an accurate representation of their realized long-term diet (Tieszen et al. 1983). In addition, prey items from gut content analysis of museum
specimens and fecal samples taken from live individuals cannot always be reliably identified and soft-bodied prey species can be underrepresented (Plummer 1991). Thus there is the need to supplement these approaches for determining diet with other methods that offer a less biased, longer-term assessment of what prey are being eaten.

An increasingly widely used method that addresses these issues is stable isotope analysis (SIA) of diet which has been increasingly used in ecological studies because it provides a more integrated estimate of diet over a six-month period (Pilgrim 2005) rather than simply a description of the last meal for a given individual. (Newsome et al. 2009; Roth et al. 2007; Vander Zanden et al. 2010; Wise et al. 2006). Many elements have at least two stable isotopes, one of which is much more common than the other in nature, and SIA measures the relative amount of each isotope thereby giving an isotope ratio (i.e. C$_{13}$:C$_{12}$) (Hobson & Wassenaar 1999). When an organism ingests prey items, the isotopic signature of the prey is incorporated into the animal’s tissue. Hence by measuring the ratios of these isotopes in various prey items (the source) as well as the predator (the mixture), one can estimate the components of different prey in a predators diet (West et al. 2006). Elements most commonly used for diet analyses in ecological studies are carbon and nitrogen (C$_{13}$:C$_{12}$, and N$_{15}$:N$_{14}$ respectively) (Gannes et al. 1997) and these elements can vary based on geography (Gustafson et al. 2007), trophic level (Urton & Hobson 2005) and vegetation type (Koch et al. 1995).

To date, mixing models have been the most common technique used to estimate the proportion of sources in a mixture, allowing the researcher to determine the unique
mathematical solution from $n+1$ sources using $n$ isotope signatures (Phillips et al. 2005). For example, Newsome et al. (2009) used these methods to estimate levels of individual diet specialization in California sea otters (Enhydra lutris nereis) and Ikeda et al. (2010) used a combination of mixing models to investigate the diet composition of 14 species of harpaline ground beetles. However, these models failed to incorporate variability in source means and could not account for variation among populations or variation among individuals within a population but simply provided an overall estimate of diet from a single population. Recently, Semmens et al. (2009) developed a Bayesian modeling framework to simultaneously estimate diet while accounting for variation at different hierarchical scales, such as among populations or individuals. This promises to allow researchers to quantitatively assess the level of variation within a given system while accounting for diet variability, which will provide a valid method for comparison of diet variation among different hierarchical levels, for example among populations or individuals (Semmens et al. 2009).

Massasauga rattlesnakes (Sistrurus catenatus) are small vipers currently grouped into three subspecies which range from the southwestern United States through southern Ontario, Canada (Campbell & Lamar 2004). Eastern massasaugas (Sistrurus c. catenatus) are distributed throughout the Midwestern United States and into Canada in a series of isolated populations occurring in a wide range of habitats (Campbell and Lamar 2004). Previous studies of eastern massasauga diet based on gut content and fecal analysis have yielded a variety of results. For example, Keenlyne and Beer (1973) have shown that small mammals, particularly voles, are the preferred prey and constitute the
majority of the diet (~ 93%). Recently, Weatherhead et al. (2009) showed that diet of eastern massasaugas in Ohio and Ontario consisted primarily of two species of shrews; similar results were shown for neonate massasaugas in Illinois (Shepard et al. 2004). Hallock (1991) studied eastern massasauga diet in Michigan and showed that mammals also constituted the majority of the prey items consumed (76% of total diet), followed by snakes (15%). However, because these studies used gut content analysis often in combination with fecal samples to estimate diet they may be subject to some of the biases described above. Furthermore, these studies fail to quantitatively compare variation in diet across populations on a range-wide scale. In addition, a significant gap in our knowledge is the degree to which these snakes show preference for specific prey relative to their availability in the environment. The diet of specific snakes has been shown to vary with both small mammal abundance (Santos et al. 2008) and prey availability (Santos et al. 2000) but it is unclear as to the general degree to which snakes prefer (or avoid) specific prey relative to their abundance (Greene 1986).

Here, I conduct SIA analyses using carbon (\(^{13}\)C) and nitrogen (\(^{15}\)N) stable isotopes to address a number of important questions regarding diet in these snakes. Specifically, first, I use a Bayesian mixing model framework to examine the relative importance of individual versus population level variation on diet variation among populations of eastern massasauga rattlesnakes. I also use this same modeling approach to investigate temporal variation in diet from a population in central Ohio. After this, I estimate the diet composition of each population separately. I then compare the results from my stable isotope analysis of diet to estimates of mammal abundance at my study sites from live
trapping data to determine if massasaugas show a preference for or avoidance of specific mammalian prey types. Finally, I present dietary results from a geographically disjunct population of western massasaugas (*Sistrurus c. tergeminus*) in Colorado, which has been shown via fecal and gut content analyses to have a distinct diet from *S. c. catenatus* (Holycross & Mackessy 2002). Together my results represent an unusually comprehensive analyses of diet composition based on SIA for a snake on a range-wide scale that addresses important questions about individual and population level variation in diet as well as whether there is evidence for preference and for which prey. These results provide important information that can be used to address ecological, conservation and evolutionary questions about these snakes.

Methods

*Sample Collection*

My overall goal was to compare the diet of this species over space (geographic variation) and time (temporal variation within one population); sampling multiple individuals from each population provided us with the framework to accomplish this goal (Fig 4.1). To investigate among population differences in diet, with the help of many collaborators, snake ventral scales were collected from massasauga rattlesnakes from each of 5 populations across the species range between the months of April and September 2007 and from one population of western massasauga in Colorado during summer 2006. Snake ventral scales have been shown to be a good tissue for diet reconstruction in rattlesnakes because over time they equilibrate to a controlled laboratory diet in
predictable ways (Pilgrim 2005). To estimate temporal variation in massasauga diet I also collected snake ventral scales from a single population (Killdeer Plains in central Ohio) for three consecutive summers from 2005-2007. Prey samples and snake samples were collected from the same season and year because isotopic compositions in specific environments can vary leading to erroneous results (Pilgrim 2005). Mammalian prey was sampled using a combination of snap traps, Sherman live traps and drift fence arrays leading into pit-fall traps. Mammal prey consisted primarily of four species, *Microtus pennsylvanicus* (meadow vole), *Peromyscus* species (white-footed or deer mouse), *Blarina brevicauda* (northern short-tailed shrew) and *Sorex cinereus* (masked shrew) with the occasional occurrence of *Napeozapus* and *Zapus* (jumping mice), *Blarina carolinensis* (Southern short-tailed shrew) and *Cryptotis parva* (Least shrew). Non-mammalian species consisted primarily of *Rana clamitans* (green frog), *Rana pipiens* (northern leopard frog), *Thamnophis sirtalis* (eastern garter snake) and *Storeria dekayi* (northern brown snake) all of which were sampled opportunistically throughout the season from each population. Mammal prey from Colorado consisted of arthropods (*Scolopendra* centipedes, Family: *Scolopendridae*), lizards (*Sceloporus undulatus*, *Holbrookia maculata*, *Eumeces obsoletus*), and various mammal species (*Peromyscus maniculatus*, *Onychomys leucogaster*, *Dipodomys ordii*, *Perognathus flavus*, *Perognathus flavescens*).

**Sample Preparation and Analysis**

To eliminate water from samples, snake scales were dried for a period of 24 hours and
individual prey items were freeze-dried. To estimate diet, I sampled prey tissues in a way that accurately reflects how snakes consume prey leading to two important steps during sample preparation. First, because mammal hair is largely undigested, the skin was removed from each mammalian prey item. Second, because snakes swallow prey whole, freeze-dried individual prey items were homogenized using a ball mill to obtain a fine powder for analysis that represents an average of tissues from the whole animal. Homogenized samples were then run through a Costech Elemental Analyzer before being processed using a Finnigan Delta Plus XL mass spectrometer at the University of Michigan Biological Station. For each sample, I estimated carbon (C\textsubscript{13}/C\textsubscript{12}) and nitrogen (N\textsubscript{15}/N\textsubscript{14}) stable isotopic ratios and results are presented in standard delta notation (\(\delta\)) based on the following equation:

\[
\delta^{13}C \text{ or } \delta^{15}N = [(R_{\text{sample}}/R_{\text{standard}}) – 1] \times 1000,
\]

where R is the ratio of C\textsubscript{13}/C\textsubscript{12} or N\textsubscript{15}/N\textsubscript{14}. The standards used were Vienna Peedee Belemnite for carbon and atmospheric N\textsubscript{2} for nitrogen and I determined instrument precision by using caffeine as a quality assurance sample for each run.

*Estimating diet from stable isotope data*

I was first interested in estimating diet for each population of massasauga so I calculated the mean and standard error for both \(^{13}\text{C}\) and \(^{15}\text{N}\) for snakes and each species of prey on a population-by-population basis. For computational efficiency, I combined biologically
similar prey species for all analyses resulting in five biologically and taxonomically
distinct groups; frogs, mice, shrews, snakes and voles for eastern massasaugas and
arthropods, lizards and three groups of mice for western massasaugas. A major
advantage of combining prey species is that it reduces the number of prey groups in the
analysis thereby reducing model complexity often leading to more robust and
biologically meaningful estimates of diet (Phillips et al. 2005). However, this is only
valid if the taxonomic groups being combined are isotopically similar. I tested for this
using ANOVA’s on prey isotope values and found no differences among the species of
frogs, shrews and snake so I was confident in my ability to combine species within in
these groups to form larger taxonomic groups (results not shown). Similarly for western
massasauga prey species, I was unable to differentiate isotope values among Onychomys
leucogaster and Dipodomys ordii - and Perognathus flavus and Perognathus flavescens.
These were clustered into groups named mouse-2 and mouse-3, respectively (e.g. see
Figure 4.2).

A critical step in using SIA to estimate diet is choosing a proper fractionation value for
each isotope (Caut et al. 2009). I first attempted to use the results from a controlled
experiment of a congeneric species, the dusky pygmy rattlesnake (Sistrurus m. miliarius)
(Pilgrim 2005) but these values (1.2‰ for $^{15}$N and 4.0‰ for $^{13}$C) resulted in mixing plots
in which massasaugas (the mixture) were located outside the isotopic polygon formed by
connecting the outermost prey species (the source) in the plot of carbon and nitrogen
mean values (as in Fig 4.2). This can happen for one of two reasons: the use of incorrect
fractionation values or because of failing to sample all potential prey species. Because
my prey sampling was based on the results of multiple previous studies of diet based on gut and fecal analyses it is unlikely I failed to sample potential prey species. Therefore, I feel this result was more likely due to incorrect fractionation values. As an alternative, I used mean trophic fractionation values determined from Post (2002) of 3.4‰ for $^{15}$N and 0.39‰ for $^{13}$C that are commonly used fractionation values for SIA studies of diet in animals.

*Hierarchical Structure of Isotopic Variation*

I use two novel approaches for estimating diet in this species. First, to determine the level of structure (individual vs. population) that explained the most variation in my samples I compared results from eight hierarchically structured Bayesian models modified from Semmens *et al.* (2009) to ask if massasauga diet varied across both space (geographic analysis) and time (temporal analysis). The analyses were run separately to test for geographic variation and temporal variation. For the geographic analysis, the following four models were tested with regard to patterns of individual and population level variation in diet; model 1: All snakes share a single similar diet with no differences among populations or individuals; model 2: Dietary differences exist among populations but individuals within populations share a similar diet; model 3: Dietary differences exist between populations as well as among individuals; model 4: Populations share a similar diet but diets among individuals within populations are variable. When assessing the fit of these models to the data, it is important to account for variation in the isotopes not attributed to diet so each of the four model types was run both with and without the use of
a residual error term. I then repeated this analysis using one population (KL) with samples collected in three consecutive years (2005 – 2007) to determine if diet was similar over time (temporal analysis). In this analysis, each year was treated as a separate population as above using the same set of models to determine the level of structure (year vs. individual) explaining the highest amount of variation in my dataset. Each model, for both analyses, was run for 300,000 iterations with a burnin of 50,000 iterations using Gibbs sampling and three parallel chains in JAGS (Plummer 2003) called from R (R Development Core Team 2010). The deviance information criterion (DIC) was used to determine the model with the best fit to the data.

*Single population estimates of diet*

Next, I used two Bayesian models to estimate diet separately for each population (assuming regional variation; see Results below). The first model (similar to the mixSIR model) treats the source means (prey means) as fixed values in the model with known mean and standard deviations. The potential problem with this approach is that often SI studies of diet suffer from small source sample sizes and therefore the calculated mean and SD can be misleading. To alleviate this problem, the second model (fully Bayesian model) treats the source means as unknown parameters in the model with their own statistical distribution (Ward et al. 2010). I ran both models and compared the results among populations. Each model consisted of three MCMC chains run for 2,000,000 iterations with a burnin of 50,000 iterations because preliminary analyses of runs this length indicated they were long enough to achieve convergence.
Prey Preference

To determine whether snakes showed either a preference or avoidance for certain mammals in the diet I calculated Manly’s alpha selectivity index (Manly et al. 1972; Randa et al. 2009) using a dataset of small mammal abundance collected throughout the state of Ohio from 2004 – 2005 (Pietkiewicz and Harder, unpubl. data). It’s important to note that this trapping occurred a few years prior to my sampling for SIA so I am assuming small mammal abundance has remained relatively constant over that time. Also, Pietkiewicz and Harder (unpubl. data) trapped in only two of my three populations (KL and GRL-1); for PRF, I use trapping data collected by Pietkiewicz and Harder (unpubl. data) from a location only 4 km from PRF and located within the same county. Furthermore, the Ohio Gap Project land cover map has both PRF and the Pietkiewicz location coded as open canopy grassland suggesting they contain very similar habitat (Ramirez et al. 2005). For a detailed description of the trapping methods used see Harder et al. (in prep). Briefly, three different trap types (pitfall, Sherman, and snap) were set along transects and mammals were identified to species level. This yielded a dataset with information on the number of individuals of each species for each transect at each site. I used this dataset rather than data collected via my trapping methods because unlike my trapping protocol with a goal to obtain as many prey as possible per unit effort, this alternate data set was specifically designed to accurately estimate prey abundance. Manly’s alpha selectivity index was then used to determine if prey items are being selected in proportion with their occurrence in the environment and was calculated as follows:
\[ \alpha_i = \frac{r_i/n_i}{\sum_{j=1}^{m} r_j/n_j} , \]

with \( r_i \) and \( r_j \) representing the number of prey types \( i \) and \( j \) in the diet, \( n_i \) and \( n_j \) representing the number of prey items \( i \) and \( j \) in the environment and \( m \) as the total number of prey species. This index accounts for the relative abundance of different prey species in the environment with \( \alpha_i \) values equal to \( 1/m \) representing a population feeding on prey in accordance to their abundance in the environment. Alpha values range from zero to one; zero represents a strong avoidance and one represents a strong preference for a given prey item. I calculated an alpha value for each prey type in each population using the mammal trapping dataset as my environmental abundance and results from the fully Bayesian model as the dietary abundance (Manly et al. 1972).

Though Manly values are informative, they are simply compared to a value of \( 1/m \) to determine if individuals are selectively choosing or avoiding prey and give no indication as to how different each calculated value is to \( 1/m \) (Manly et al. 1972). For example, a calculated value of 0.25 would suggest a population is avoiding a given prey type but further statistical analyses are needed to determine if this value is less than expected through chance alone. To estimate the significance of each alpha value, I used R (R Development Core Team 2010) to simulate 10,000 random alpha values from the observed number of prey items \( n \) in the wild. For each population, I used a bootstrapping method to resample from the known prey abundance in the wild. This created 10,000 random snake diets, each of which consumed \( n \) prey items. From these
simulated snakes, I then calculated 10,000 $a_i$ values for each prey item. If snakes are taking prey items relative to their abundance in the wild, I expect alpha values of $\sim 0.33$ ($a_i = 1/m$) because I have three potential prey items so I report p-values as the proportion of simulated alpha values either greater than or less than 0.33.

Results

Snake and Prey Isotope Values

I collected a total of 96 snake scale samples from five populations of the eastern massasauga (range: 9 – 29 per population) and 16 scale samples from the western massasauga (Table A.5). Prey sample sizes averaged 10.2 individuals per species per population (Table A.5). Average $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values from snake scales showed significant variation among populations (ANOVA: $\delta^{13}\text{C}: F_{4,92} = 11.95, p < 0.001; \delta^{15}\text{N}: F_{4,88} = 30.30, p < 0.001$) (Fig. 4.2). Relative to their prey, massasaugas were more enriched in $^{15}\text{N}$ in all eastern massasauga populations, which is expected given typical trophic enrichment often observed in this isotope (Figure 4.2). Prey isotopic values showed some general patterns across populations. In four of five populations, voles were most depleted in $^{13}\text{C}$ and mice were most enriched in $^{13}\text{C}$. Similarly, shrews and frogs were most often most enriched and voles and mice were most depleted in $^{15}\text{N}$. Scatterplots of mean $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ for all prey items revealed significant differences between nearly all prey items in each population. The exceptions were frogs and shrews in KL, PRF, SSSP and CS (Fig. 4.2).
Hierarchical Structure of Isotopic Variation

My hierarchical modeling approach for temporal variation showed that model 2, with a residual error term but no temporal or individual variation terms, fit the data best (DIC = 328.3). Model 4, which included a residual error term and a temporal variation term but no individual variation term, was the second best fitting model but had a DIC value (331.5) that was significantly lower than the value for model 2. In general, models that included a residual error term had lower DIC values suggesting the incorporation of an error term improves model fit. There was little support for models that did not include a residual error term (models 1, 3, 5, 7) compared to models including a residual error term (models 2, 4, 6, 8).

For the geographic models, model 4, which included a population variation term and no individual variation term, fit the data best (DIC = 658; Table 4.1). Model 2 containing a residual error term but no population variation term was the second best model (model 2: DIC = 660.3; Table 4.1). Models with individual variation terms were among the least favored models (models 5 - 8). Overall, these results suggest isotopic variation in this dataset is best explained by a model which that incorporates differences in diet among populations but with no variation in diet between individuals within populations.
Single population estimates of diet

For single population estimates of diet, results from both the fully Bayesian model and the mixSIR model were similar although, as expected, the 95% confidence intervals for each prey type were wider for the fully Bayesian model (Ward et al. 2010). Results from the fully Bayesian model for each population showed that mammals were the most commonly-consumed prey category in four of five populations (Fig. 4.3a) with shrews (Blarina and Sorex) the most common prey item in the diet in three of these four populations (50 – 82%). Results from KL suggested frogs were the most common prey item taken. However, due to the results of my temporal analysis (see below) and my inability to differentiate between frogs and shrews at KL, I believed this to be an analytical artifact. To test this, I subsequently removed frogs from the analysis and ran the model again. At this point, shrews became the most dominant prey item (82%) with all other prey items remaining at levels similar to the original model run. In light of this result, I thereafter pooled frogs and shrews into one group and ran the model once more for this population. The new frog-shrew group comprised 82% of the diet composition in this population (results not shown). Although I could not detect a difference between isotope values for frogs and shrews at KL results from previous studies using gut contents and fecal samples (Weatherhead et al. 2010) suggests that shrews are the more likely prey item of the two.

Results from Colorado for the western massasauga rattlesnake suggest insects (33%) and lizards (17%) are the most common prey items rather than one of three separate types of mammals (Figure 4.4). However, when the three mammal groups are combined into a
single prey category, they become the second most important prey item, representing 23% of the diet.

**Mammal Abundance and Prey Preference**

Due to the availability of abundance data I was able to assess whether eastern massasauga rattlesnakes showed evidence for population-level preferences for particular mammal species. Four species comprised the majority of the mammals trapped at each site (*Microtus pennsylvanicus, Peromyscus spp., Sorex cinereus,* and *Blarina brevicauda: Sorex* and *Blarina* were combined into “shrews”). The Manly alpha index revealed a preference for shrews (2 species combined) in two of three populations (KL and PRF) while snakes avoided them in the third (GRL-2) (Fig. 4.5, Table 4.2). Snakes showed a preference for Peromyscus at GRL-2, no preference at PRF, and avoided them at KL (Fig. 4.5, Table 4.2). Finally, there was a strong avoidance of voles in all three populations (Fig. 4.5, Table 4.2).

**Discussion**

Many studies have used stable isotope techniques to estimate diet often by using mixing models (Moore & Semmens 2008). Recently, there have been significant improvements in the methods used to estimate diet by implementing Bayesian statistics that provide the ability to simultaneously estimate diet across time and space (Semmens et al. 2009) and improve dietary estimates when limited by small source sample size (Ward et al. 2010). This is especially important in studies like ours because it is often difficult to obtain large
numbers of prey samples for each species. These fully Bayesian methods derive dietary estimates by treating source means (prey means) as model parameters rather than fixed, known estimates and this has been shown to lead to less biased estimates of diet (Ward et al. 2010). Here, I use these novel approaches to shed light on the patterns of intraspecific variation in massasauga rattlesnake diet. Results suggest diet varies geographically among populations whereas individuals within each population share similar overall diet. Furthermore, there is no evidence of temporal variation over a three-year time period at a population in central Ohio. Massasaugas display a preference for shrews in two of three populations in Ohio and an avoidance of voles in all three populations. Below, I discuss the implications of these results.

SIA versus Conventional Methods of Estimating Diet

My results suggest mammals are the primary prey species in three of five populations. This is consistent with past studies based on gut and fecal analyses demonstrating that eastern massasauga diet consists primarily of mammalian prey including Blarina, Microtus, Peromyscus, Sorex, Tamias and Zapus with frogs and small snakes making up the remainder of prey consumed (Keenlyne & Beer 1973; Weatherhead et al. 2009). Based on SIA, western massasaugas from Colorado display a mixed diet consisting of arthropods (centipedes: 33%), lizards (18%) and mice (22%). This is largely consistent with previous studies of gut content from snakes within the same population obtained from field encounters and museum specimens (Holycross & Mackessy 2002). They found lizards to be the most prominent prey species representing 54.4% of snake diet
followed by mammals (40.4%) and arthropods (centipedes: 5.3%) but perhaps more importantly, SIA revealed a similar mixed diet when compared to gut content analysis with the exception of centipedes being more important than gut content analysis suggested. When compared to the diet of the eastern massasauga, western massasaugas consume more arthropods and lizards (ectothermic prey) than eastern massasaugas (primarily mammal based diet).

*Patterns of Variation in Diet – Geographic Variation*

Geographic variation in diet has been documented in other snake species with genetic differences in prey preference (Arnold 1981), and differences in prey availability and abundance (Santos et al. 2008) being cited as possible mechanisms. To date, there have been few studies of massasauga diet that have incorporated comparisons among populations and subspecies (Holycross & Mackessy 2002; Weatherhead et al. 2009) compared to those that describe diet in a single population (Ditmars 1908; Hallock 1991; Keenlyne & Beer 1973; Pope 1926; Shepard et al. 2004) and this study is the first to compare diet across the range of this species using SIA. In GRL-2, *Microtus* comprised the bulk of the diet (50%) while snakes from KL, CS and SSSP fed primarily on shrews. *Microtus* appears less important to the overall feeding ecology of massasaugas when compared to *Sorex* and *Blarina* and is somewhat surprising given the previous studies that report an abundance of *Microtus* in the diet (Keenlyne & Beer 1973). But, juvenile massasauga diet was studied at SSSP using gut content analysis (Shepard et al. 2004) where the bulk of the diet consisted of both shrews (*Blarina carolinensis*) and voles.
(*Microtus ochrogaster*) and my results from SIA corroborate these findings; SIA of diet at SSSP yielded 56% shrews and 20% voles.

My results support findings from previous studies that suggest frogs and small snakes in the diet is geographically variable and may be determined by differences in habitat and/or abundance of ectothermic prey (Shepard et al. 2004). When comparing populations, frogs are important prey items in both PRF and potentially KL whereas they are less important in CS and GRL-2. The inclusion of frogs in the diet may be a direct result of their abundance in a particular population, as previously authors have suggested frogs are among the most important prey items in specific populations (Ditmars 1908; Pope 1926). Indirect evidence for this comes from a study by Reiserer (2002) in which geographic variation in caudal luring (behavior used to capture frogs) was documented within several *Sistrurus* species providing further evidence that local frog abundance may be an important factor giving rise to the incidence of frogs in massasauga diet.

Variation in diet among populations can help explain previously documented variation in two key phenotypic traits in these snakes: body condition and venom protein composition. First, in Chapter 3 I document significant differences in body condition among populations. Though not directly tested here, it remains possible that geographic differences in diet may help explain the pattern of variation in body condition among populations. A poor diet could lead to malnourished individuals and hence have a large impact on individual body condition. Poor body condition has been shown to have dramatic effects on other species of snakes by reducing female reproductive ability
and survival rates (Madsen & Shine 1993), both of which can lead to population size reduction and lower overall fitness for a population.

Geographic variation in diet may also have important implications for the evolution of venom proteins in these snakes. A significant level of geographic variation in venom proteins has been documented in this species yet the reasons for this are currently unknown (Gibbs and Chiucchi 2011). Many authors have hypothesized that local adaptation to differences in available prey species would create a pattern whereby population-level differences in venom proteins were reflective of overall variation in diet (Daltry et al. 1996). Although I couldn't directly test this hypothesis, my results show population level differences in diet among the same populations that exhibit population differentiation in venom proteins and are therefore consistent with the above hypothesis.

*Patterns of Variation in Diet – Individual Variation*

An advantage of using SIA to study diet is that it can provide insight into hierarchical levels of variation at the same time. For example, the Bayesian modeling framework implemented here allowed us to quantify levels of individual variation and population variation using a comparative modeling approach.

I see no evidence for individual diet variation in *S. catenatus* within a population. A model without an individual variation component was favored over any model with an individual variation component. This result suggests individuals within each population
share similar overall diets with an absence of individual diet specialization in \textit{S. catenatus}. To the best of my knowledge, individual specialization in diet in snakes has not been documented previously and a lack of individual specialization in snakes has previously been documented only once; in a garter snake population using mark-recapture and gut-content analysis (Kephart & Arnold 1982). Total population niche width (TNW) is the sum of within-individual variation component (WIC) in diet and between-individual variation component (BIC) in diet. Low levels of individual variation in diet in \textit{S. catenatus} suggest each population’s WIC is larger than its BIC (Bolnick \textit{et al.} 2003). This has important conservation implications because populations with low levels of BIC and hence little individual specialization are often less stable with reduced ability to adapt in a changing environment (Durell 2000). But, these populations may prove easier to conserve through management practices because it allows managers to target conservation efforts toward preservation of the “average” prey resource base in each population because individuals do not specialize on just one type of prey but rather can utilize a wider range of available resources (Bolnick \textit{et al.} 2003).

\textit{Patterns of Variation in Diet – Temporal Variation}

To my knowledge, this represents one of the first studies of snakes to show stable diets over multiple years. Willson \textit{et al.} (2010) analyzed annual and seasonal variation in stable isotopes of two sympatric snake species, the banded watersnake (\textit{Nerodia fasciata}) and the black swamp snake (\textit{Seminatrix pygaea}). Their results suggest diet stability may vary with species; prey consumed by the banded watersnake changed in accordance with
changes in the prey base but remained relatively stable over time in the black swamp snake. Here, my results suggest massasauga diet has remained relatively stable over a much longer period (three years with shrews being the primary prey item in each of the three years). Given that small mammal abundance has changed little over this time period in the state of Ohio (Pietkiewicz unpubl. data), this result is not surprising. However, I caution that three years likely isn’t long enough to pick up on any small mammal cycles that might occur at this site and thus may not be representative of snake diet over a much longer time period. In contrast, given that shrews appear to be a preferred prey choice for this species, it is possible that fluctuations in mammalian abundance would not cause a shift in diet as long as shrews remain a viable option for snakes.

Prey Preference

In terms of prey preference in snakes, Shine (1991) suggests snake species with diverse dietary habits should exhibit geographic variation in diet that closely resembles local prey diversity and abundance and this pattern has been shown in other snake species. For example, Daltry et al. (1998) showed dietary patterns of the Malayan pit viper (Calloselasma rhodostoma) closely resemble the pattern of local prey abundance. This hypothesis is further supported by evidence from terrestrial garter snakes (Thamnophis elegans) in which snake diet from 22 locations closely mirrored the relative abundance of local prey species (Burghardt 1993; Kephart 1982). Here, I find evidence that based on SIA massasaugas are selectively choosing shrews over other mammalian prey species.
that are more abundant suggesting prey availability and abundance may not be the determining factor for local diet in this species. Rather, because massasaugas are feeding on mammals out of proportion with mammal abundance, genetic differences in prey preference due to divergent selection may be responsible for the observed preferences, as previous studies have documented prey preference is a heritable trait in snakes (Arnold 1981; Burghardt 1993).

The fact that shrews are the most prominent prey item in three of five populations and snakes show a preference for shrews in two populations adds to the growing literature suggesting shrews are an important prey item in many populations across the range of this species (Holycross & Mackessy 2002; Shepard et al. 2004; Weatherhead et al. 2009). From a snakes’ perspective, shrews may be preferred over other mammalian prey for two primary reasons. First, although shrews represent a much smaller meal, they are more active than other mammals known to have a much higher metabolic rate (Platt 1974; Tomasi 1984). This could cause the rate at which snakes encounter shrews to be much higher than the encounter rate for voles and mice despite the lower abundance of shrews in many populations. Second, massasauga venom may be more efficient against smaller species with a high metabolism by increasing the speed with which the prey item is incapacitated thereby limiting the distance a given prey item can travel after injection of venom and thus increasing the likelihood that the snake can find this prey item and ingest it.

Of equal interest is the apparent avoidance of *Microtus* (voles) in all three populations.
This result is somewhat surprising given the high occurrence of voles in the diet in certain populations (Keenlyne and Beer 1973). One possibility is that many voles are simply too large for all but the largest massasaugas to ingest. While this seems plausible, snakes are capable of ingesting prey with relative mass ratios (MR, ratio of prey mass to predator mass) much greater than one (Cundall & Greene 2000; Greene 1983; Pough & Groves 1983; Shepard et al. 2004). The average mass of voles captured at Killdeer Plains during my study was 32.6 grams compared to the average mass for adult massasaugas in my study of 291.5 grams resulting in a small MR of 0.11. This value is much smaller than MR’s reported for neonates of this species in Illinois (Shepard et al. 2004) (average MR = 0.47) so it seems unlikely that prey size limits ingestion of Microtus in this species. Rather, a more likely explanation can be considered given the natural history of Microtus species and their preference to use runways as methods of travel through habitats generally consisting of dense vegetation (DeCoursey Jr 1957). Massasaugas are generally considered sit-and-wait predators and unless they choose ambush sites along runways, their likelihood of encountering voles may be much lower than the likelihood of encountering shrews. Whether the observed preferences for shrews and avoidance of voles in my study is a result of divergent selection in isolated populations remains to be seen and future work should focus on testing this hypothesis.

Analytical Issues

Here, I have demonstrated differences among populations in diet based on SIA but it is important to recognize that these results are based on a number of important assumptions
affiliated with the design of my study and should be interpreted with some caution. SIA of diet relies on choosing proper fractionation values for the species being studied and most often these values are unknown. This is true for eastern massasaugas as well though there has been some work done on a congeneric species, the dusky pygmy rattlesnake (*Sistrurus m. miliarius*) (Pilgrim 2005). When running my diet models, I first used the fractionation values established in captivity by Pilgrim (2005) however these values routinely yielded inconsistent results that failed to converge. Further, after correction using the values from Pilgrim (2005), massasauga mean values for $^{15}$N were outside the isotopic polygon (Fig 4.2) in all populations suggesting incorrect fractionation values. Furthermore, the fractionation values of Pilgrim (2005) are quite different than those often reported for carnivorous vertebrates (Post 2002) with high protein diets. Therefore, I used average values from the literature summarized in Post (2002) and with these values massasauga mean $^{15}$N values consistently fell within the isotopic polygon for each population. The choice of fractionation values is an important one because it can directly affect the results of dietary analyses so caution should be used when interpreting my results. That being said, I varied fractionation values by 0.5 (delta units) and found this had little effect on the results and major conclusions were unaltered. Because of this, I am confident in my major conclusions though I do caution against interpreting specific values for each dietary category and rather suggest defining diet based on prey species distributions from the fully Bayesian model.


Table 4.1: Summary of results from the Bayesian hierarchical modeling approach modified from Semmens et al. (2009) showing the model numbers and specific terms included in each model. A value of “N” indicates the term was not included in a given model and a value of “Y” indicates the term was included in the model. DIC is the deviance information criterion of model selection. Smaller numbers represent better overall model fit and the box highlights the best model.

<table>
<thead>
<tr>
<th>Model</th>
<th>Regional Variation</th>
<th>Individual Variation</th>
<th>Residual Error Term</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>1189.3</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>660.3</td>
</tr>
<tr>
<td>3</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>836.1</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>658.0</td>
</tr>
<tr>
<td>5</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>1027.0</td>
</tr>
<tr>
<td>6</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>661.9</td>
</tr>
<tr>
<td>7</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>911.7</td>
</tr>
<tr>
<td>8</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>668.3</td>
</tr>
</tbody>
</table>
Table 4.2: Results of Manly’s alpha indices calculated for each population and each prey type. Values greater than 0.33 suggest a preference for a given prey type whereas values less than 0.33 suggest avoidance of a given prey type. Values in parentheses represent the 95% confidence interval from 10,000 Manly values randomly generated in R (R Development Core Team 2010).

<table>
<thead>
<tr>
<th></th>
<th>KL Manly Index</th>
<th>p-value</th>
<th>PRF Manly Index</th>
<th>p-value</th>
<th>GRL-2 Manly Index</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shrew</td>
<td>0.79 (0.67-0.89)</td>
<td>0.001</td>
<td>0.52 (0.31-0.70)</td>
<td>0.020</td>
<td>0.14 (0.05-0.22)</td>
<td>0.001</td>
</tr>
<tr>
<td>Mouse</td>
<td>0.15 (0.05-0.28)</td>
<td>0.017</td>
<td>0.28 (0.10-0.48)</td>
<td>0.280</td>
<td>0.67 (0.48-0.89)</td>
<td>0.001</td>
</tr>
<tr>
<td>Vole</td>
<td>0.06 (0.03-0.08)</td>
<td>0.001</td>
<td>0.20 (0.10-0.29)</td>
<td>0.007</td>
<td>0.19 (0.06-0.30)</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Figure 4.1: A) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. State abbreviations are as follows: IL (Illinois); IN (Indiana); KY (Kentucky); OH (Ohio); WV (West Virginia); PA (Pennsylvania); NY (New York); MI (Michigan); ON (Ontario, Canada). B) Inset showing the relative location of the sampling locations within the United States and the location of the western massasauga locality in Colorado (CO).
Figure 4.2: Mean and standard error of carbon ($\delta^{13}C$) nitrogen ($\delta^{15}N$) isotopic composition prior to any correction for fractionation for each prey taxa and massasauga rattlesnakes. Results are shown for each of five eastern massasauga rattlesnake populations (ILL, KL, PRF, GRL-2, CS) and one western massasauga rattlesnake population (CO). Massasauga mean values are displayed within a box to highlight their position relative to potential prey items.
Figure 4.3: Distribution of prey items in the diet of A) eastern massasauga rattlesnakes determined using the fully Bayesian model for each population independently. Each column represents specific prey taxa while each row represents a given population. Median values are reported above each individual graph and the taxa representing the highest proportion of the diet is marked with a box.
Figure 4.4: Distribution of prey items in the diet of western massasauga rattlesnakes determined using the fully Bayesian model. Each graph represents a specific prey taxa. Median values are reported above each individual graph and the taxa representing the highest proportion of the diet is marked with a box.
Figure 4.5: Comparison of the abundance of mammalian species as determined by mammal trapping at each site (Piekiewticz unpubl. data) (upper graph) and the percentage of each mammalian prey type as calculated using SIA under the fully Bayesian model (lower graph) from each of three populations.


Coltman DW, Bowen WD, Wright JM (1998) Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. Proceedings of the Royal Society B-Biological Sciences 265, 803-809.


Luiselli L, Agrimi U (1991) Composition and variation of the diet of *Vipera aspis francisciredi* in relation to age and reproductive stage. *Amphibia- Reptilia* 12, 137-144.


Table A.1. Summary of 19 polymorphic microsatellite loci. Null alleles were estimated using FreeNa. * represents $F_{IS}$ values significant after correction using the B-Y method (Narum 2006). Number of alleles was calculated from 388 individuals.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Number of alleles</th>
<th>$H_o$</th>
<th>$H_e$</th>
<th>$F_{IS}$</th>
<th>Frequency of Null Allele(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scu 200</td>
<td>29</td>
<td>0.75</td>
<td>0.80</td>
<td>0.08*</td>
<td>0.03</td>
</tr>
<tr>
<td>Scu 201</td>
<td>12</td>
<td>0.69</td>
<td>0.72</td>
<td>0.06</td>
<td>0.02</td>
</tr>
<tr>
<td>Scu 202</td>
<td>15</td>
<td>0.59</td>
<td>0.65</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Scu 203</td>
<td>7</td>
<td>0.48</td>
<td>0.50</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Scu 204</td>
<td>9</td>
<td>0.45</td>
<td>0.50</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>Scu 205</td>
<td>11</td>
<td>0.72</td>
<td>0.72</td>
<td>-0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Scu 206</td>
<td>4</td>
<td>0.16</td>
<td>0.17</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Scu 207</td>
<td>11</td>
<td>0.46</td>
<td>0.46</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Scu 208</td>
<td>9</td>
<td>0.61</td>
<td>0.59</td>
<td>-0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Scu 209</td>
<td>8</td>
<td>0.38</td>
<td>0.38</td>
<td>-0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Scu 210</td>
<td>19</td>
<td>0.63</td>
<td>0.65</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Scu 211</td>
<td>13</td>
<td>0.74</td>
<td>0.74</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Scu 212</td>
<td>15</td>
<td>0.76</td>
<td>0.76</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Scu 213</td>
<td>18</td>
<td>0.73</td>
<td>0.73</td>
<td>-0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Scu 214</td>
<td>14</td>
<td>0.73</td>
<td>0.70</td>
<td>-0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>Scu 215</td>
<td>18</td>
<td>0.49</td>
<td>0.61</td>
<td>0.18*</td>
<td>0.09</td>
</tr>
<tr>
<td>Scu 216</td>
<td>29</td>
<td>0.64</td>
<td>0.79</td>
<td>0.15*</td>
<td>0.08</td>
</tr>
<tr>
<td>Scu 217</td>
<td>12</td>
<td>0.52</td>
<td>0.60</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>Scu 218</td>
<td>9</td>
<td>0.64</td>
<td>0.65</td>
<td>0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Table A.2. Genetic estimates of pairwise $F_{ST}$ between pairs of populations calculated using MSA based on 10,000 permutations. Values with symbols represent non-significant comparisons after Bonferroni correction.

|     | EHSP | SSSP | PRF | KL-1 | KL-2 | WM  | GRL-1 | GRL-2 | GRL-3 | JEEC | SGL95 | VEN  | BS   | BPNP | GBINP | KPP  | OJIB | WAIN |
|-----|------|------|-----|------|------|-----|-------|-------|-------|------|-------|------|------|------|------|------|------|------|------|
| EHSP| 0.10 | 0.21 | 0.11| 0.12 | 0.15 | 0.22| 0.30  | 0.21  | 0.14  | 0.18 | 0.23 | 0.27 | 0.24 | 0.14 | 0.18 | 0.20 | 0.17 | 0.17 |
| SSSP| 0.26 | 0.15 | 0.19 | 0.27 | 0.33 | 0.25| 0.24  | 0.23  | 0.26 | 0.31 | 0.28 | 0.17 | 0.24 | 0.24 | 0.23 | 0.20 |
| PRF | 0.18 | 0.21 | 0.22 | 0.30 | 0.38 | 0.28| 0.25  | 0.25  | 0.29 | 0.36 | 0.32 | 0.21 | 0.22 | 0.26 | 0.24 | 0.24 |
| KL-1| 0.01 | 0.10 | 0.17 | 0.19 | 0.13 | 0.12| 0.12  | 0.16  | 0.16 | 0.17 | 0.08 | 0.11 | 0.14 | 0.11 | 0.10 |
| KL-2| 0.12 | 0.18 | 0.20 | 0.13 | 0.13 | 0.13| 0.17  | 0.21  | 0.18 | 0.09 | 0.14 | 0.15 | 0.11 | 0.10 |
| WM  | 0.26 | 0.24 | 0.17 | 0.14 | 0.19 | 0.19| 0.27  | 0.26  | 0.13 | 0.18 | 0.18 | 0.11 | 0.14 |
| GRL-1| 0.31 | 0.16 | 0.23 | 0.24 | 0.30 | 0.35 | 0.33  | 0.21  | 0.28 | 0.22 | 0.16 | 0.21 |
| GRL-2| 0.23 | 0.29 | 0.33 | 0.29 | 0.35 | 0.32 | 0.26  | 0.31 | 0.31 | 0.27 | 0.25 |
| GRL-3| 0.19 | 0.21 | 0.24 | 0.32 | 0.28 | 0.19 | 0.25  | 0.22 | 0.14 | 0.18 |
| JEEC | 0.18 | 0.23 | 0.31 | 0.26 | 0.17 | 0.20 | 0.24  | 0.17 | 0.16 |
| SGL95| 0.22 | 0.31 | 0.32 | 0.19 | 0.23 | 0.24 | 0.18  | 0.20 |
| VEN  | 0.30 | 0.30 | 0.20 | 0.27 | 0.27 | 0.21 | 0.20  | 0.20 |
| BS   | 0.27 | 0.23 | 0.25 | 0.28 | 0.29 | 0.24 |
| CS   | 0.24 | 0.21 | 0.26 | 0.25 | 0.24 |
| BPNP | 0.12 | 0.14 | 0.16 | 0.10 |
| GBINP| 0.18 | 0.19 | 0.17 |
| KPP  | 0.17 | 0.19 |
| OJIB | 0.14 |
Table A.3. Summary of data collected across years and season for each population. The number of males and females captured from each population is also shown. The total for each column is shown at the bottom.

<table>
<thead>
<tr>
<th>Population</th>
<th>Year</th>
<th>1993</th>
<th>1994</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHSP</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>2</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>SSSP</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>2</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>PRF</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7</td>
<td>1</td>
<td>11</td>
<td>8</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>15</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>KL</td>
<td></td>
<td>-</td>
<td>8</td>
<td>16</td>
<td>9</td>
<td>-</td>
<td>13</td>
<td>30</td>
<td>9</td>
<td>7</td>
<td>13</td>
<td>33</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>WM</td>
<td></td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>-</td>
<td>11</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>GRL-1</td>
<td></td>
<td>-</td>
<td>-</td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>GRL-2</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>7</td>
<td>8</td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>GRL-3</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>6</td>
<td>11</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>PA</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>14</td>
<td>-</td>
<td>14</td>
<td>6</td>
<td>8</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>BS</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8</td>
<td>10</td>
<td>-</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>12</td>
<td></td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>CS</td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>23</td>
<td>8</td>
<td>1</td>
<td>21</td>
<td>9</td>
<td>10</td>
<td>21</td>
<td></td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>BPNP</td>
<td></td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>6</td>
<td>7</td>
<td>14</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>KPP</td>
<td></td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>10</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>WAIN</td>
<td></td>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>15</td>
<td>10</td>
<td>8</td>
<td>44</td>
<td>45</td>
<td>68</td>
<td>63</td>
<td>80</td>
<td>110</td>
<td>63</td>
<td>98</td>
<td>150</td>
<td></td>
</tr>
</tbody>
</table>
Table A.4: List of zoos contributing data for this study with corresponding number of samples from each captive population. Animals in this table were combined to form the “zoo population” for comparison to populations of wild snakes.

<table>
<thead>
<tr>
<th>Zoo</th>
<th>Number of Animals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columbus Zoo and Aquarium</td>
<td>3</td>
</tr>
<tr>
<td>Seneca Park Zoo</td>
<td>3</td>
</tr>
<tr>
<td>Detroit Zoo</td>
<td>10</td>
</tr>
<tr>
<td>Buffalo Zoo</td>
<td>2</td>
</tr>
<tr>
<td>Racine Zoological Gardens</td>
<td>1</td>
</tr>
<tr>
<td>Oklahoma City Zoo</td>
<td>2</td>
</tr>
</tbody>
</table>
Table A.5: A) Number of individual prey items and rattlesnakes captured in each of the five eastern populations. B) Number of prey items and rattlesnakes from the western population in Colorado.

### A

<table>
<thead>
<tr>
<th>Taxonomic Group</th>
<th>ILL</th>
<th>PRF</th>
<th>KL</th>
<th>GRL-2</th>
<th>CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frog</td>
<td>4</td>
<td>3</td>
<td>23</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Mouse</td>
<td>11</td>
<td>17</td>
<td>23</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Shrew</td>
<td>4</td>
<td>22</td>
<td>14</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Snake</td>
<td>8</td>
<td>17</td>
<td>11</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>Vole</td>
<td>8</td>
<td>5</td>
<td>14</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Eastern Massasaugas</td>
<td>20</td>
<td>19</td>
<td>19</td>
<td>9</td>
<td>29</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th>Taxonomic Group</th>
<th>CO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scolopendra</td>
<td>4</td>
</tr>
<tr>
<td>Lizards</td>
<td>9</td>
</tr>
<tr>
<td>Mouse</td>
<td>5</td>
</tr>
<tr>
<td>Mouse 2</td>
<td>3</td>
</tr>
<tr>
<td>Mouse 3</td>
<td>5</td>
</tr>
<tr>
<td>Western Massasaugas</td>
<td>16</td>
</tr>
</tbody>
</table>
Figure A.1: $\Delta K$ calculated with the Evanno method (2005) from the range-wide STRUCTURE analysis. There is strong support for a $K$ of 11 as determined by the relative height of the peak at this value.