MOLECULAR ECOLOGY

Molecular Ecology (2010) **19**, 5345–5358 doi: 10.1111/j.1365-294X.2010.04860.x

Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake

JAMES E. CHIUCCHI* and H. L. GIBBS*†

*Department of Evolution, Ecology and Organismal Biology, Ohio State University, 300 Aronoff Laboratory, 318 West 12th Avenue, Columbus, OH 43210-1293, USA, †Mathematical Biosciences Institute, Ohio State University, Jennings Hall, 1735 Neil Avenue, Columbus, OH 43210-1293, USA

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 because of habitat fragmentation and destruction. Here, we use data from 19 species-specific microsatellite DNA loci to assess the 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 $\theta_{Fst} = 0.21$) and a Bayesian clustering method indicates that each population represents a unique genetic cluster even at regional spatial scales. There is a twofold 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 <7 km apart. A test of alternate models of population history strongly favours 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.

Keywords: eastern massasauga rattlesnake, effective population size, historical and contemporary gene flow, microsatellite DNA loci, population differentiation

Received 17 may 2010; revision received 11 August 2010; accepted 16 August 2010

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; Schwartz & Karl 2005; Redeker *et al.* 2006). 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

Correspondence: James Chiucchi, Fax: 614 292 2030;
E-mail: chiucchi.1@gmail.com

contemporary populations in fragmented habitats may be coincidental rather than because of the direct effects of fragmentation (Cunningham & Moritz 1998; Vandergast *et al.* 2007; Hansen *et al.* 2009; Pavlacky *et al.* 2009).

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, we adopt another approach that of analysing a data set with two analytical techniques that extract information on evolutionary processes operating at distinct timescales (e.g. Hansen *et al.* 2009). These timescales bracket the period during which anthropogenic effects began to affect the landscape. Specifically, we compare assignment tests versus coalescent techniques to estimate migration rates over few versus many generations, respectively (Beerli & Felsenstein 2001; Wilson & Rannala 2003). In addition, we use measures of genetic variation to detect changes in population size over different timescales (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.* 2010). 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 subspecies for endangered or threatened status within the United States (Szymanski 1998), and it is also classified as a federal Species at Risk in Canada (Government of Canada 2009).

Here, we 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 is limited (Kubatko *et al.* 2010). Our specific goals were to (i) Characterize the overall patterns of genetic diversity as well as regional and range-wide genetic structure of eastern massasauga populations; (ii) Look for genetic signatures of population size changes; (iii) Investigate the levels of recent and historical gene flow among populations; (iv) Test among alternate models of demographic history namely migration-drift equilibrium or fragmentation followed by drift with no gene flow (Ciofi *et al.* 1999). We 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

We collected samples from adult massasauga rattlesnakes (N = 388) from 19 localities (n = 6 – 68 samples per site) throughout their range (Table 1, Fig. 1a, b). 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 2 μM 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

We used FSTAT v2.9.3 (Goudet 2001) to determine the number of alleles, observed (H_O) and expected heterozygosity (H_E), and F_{IS} for each locus. For each population, MSA v4.05 (Dieringer & Schlötterer 2003) was used to calculate allelic richness (*AR*) on a minimum of 6 individuals. Departures from Hardy–Weinberg

Table 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

Regional group	Population	ID	N	H_O	H_E	*AR*	*F*
Southern Illinois	South Shore State Park	SSSP	18	0.62	0.65	3.91	0.15 (0.12–0.19)
	Eldon Hazlet State Park	EHSP	14	0.65	0.69	4.66	0.25 (0.20–0.29)
Western and Central Ohio	Prairie Road Fen	PRF	21	0.50	0.56	3.31	0.31 (0.25–0.36)
	Killdeer Plains Wildlife Area 1	KL-1	68	0.75	0.77	5.23	0.07 (0.05–0.08)
	Killdeer Plains Wildlife Area 2	KL-2	17	0.75	0.76	4.96	0.09 (0.06–0.10)
	Willard Marsh Wildlife Area	WM	15	0.65	0.68	4.68	0.13 (0.10–0.16)
Northeast Ohio	Grand River Lowlands 1	GRL-1	18	0.58	0.57	3.31	0.35 (0.31–0.42)
	Grand River Lowlands 2	GRL-2	18	0.54	0.51	2.90	0.40 (0.35–0.49)
	Grand River Lowlands 3	GRL-3	20	0.64	0.63	3.78	0.25 (0.21–0.30)
Western Pennsylvania	State Game Lands 95	SGL95	6	0.53	0.58	3.16	0.35 (0.27–0.41)
	Venango County	VEN	7	0.50	0.57	3.80	0.23 (0.20–0.31)
	Jennings Environmental Education Center	JEEC	9	0.53	0.62	3.79	0.21 (0.17–0.28)
New York	Bergen Swamp	BS	20	0.50	0.49	2.86	0.41 (0.36–0.50)
	Cicero Swamp	CS	62	0.51	0.52	3.08	0.33 (0.28–0.37)
Ontario	Ojibway Prairie	OJIB	8	0.48	0.60	3.91	0.20 (0.14–0.23)
	Wainfleet Bog	WAIN	12	0.61	0.66	4.24	0.15 (0.11–0.19)
	Warder Ranch, Bruce Peninsula National Park	BPNP	20	0.66	0.68	4.84	0.08 (0.06–0.11)
	Beausoleil Island, Georgian Bay Islands National Park	GBINP	15	0.58	0.62	4.07	0.18 (0.14–0.22)
	Killbear Provincial Park	KPP	20	0.61	0.60	3.71	0.23 (0.20–0.29)

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 (Benjamini & Yekutieli 2001; Narum 2006). We then used FreeNA (Chapuis & Estoup 2007) to estimate a null allele frequency at each locus in each population.

We used two approaches to assess genetic differentiation among sample locations. First, we used MSA v4.05 to estimate and test the significance of Weir & Cockerham's (1984) unbiased F_{ST} estimator θ_{Fst} across all populations as well as for all pairwise combinations of populations. Second, we 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 our data. Because of marked heterogeneity in the distances between our sampling localities (Fig. 1a, b), we conducted STRUCTURE analyses at two different spatial scales. First, we conducted a large-scale range-wide analysis among populations separated by more than 50 km (see Fig. 1a). For regional groups with multiple populations within 50 km of each other (NE Ohio, central Ohio, western Pennsylvania and southern Illinois), we 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 we next removed populations containing fewer than 10 individuals from the data set. This resulted in a second range-wide data set of 11 populations which was analysed as described in the following paragraphs. Second, we also conducted separate regional analyses among each of four sets of populations located <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 Markov chain Monte Carlo (MCMC) iteration followed by 300 000 iterations using the admixture model with sampling localities as priors and correlated allele frequencies. The plot of (ln $\Pr(X \mid K)$ 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. We 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, we 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). We 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, we used the programs BAYESASS

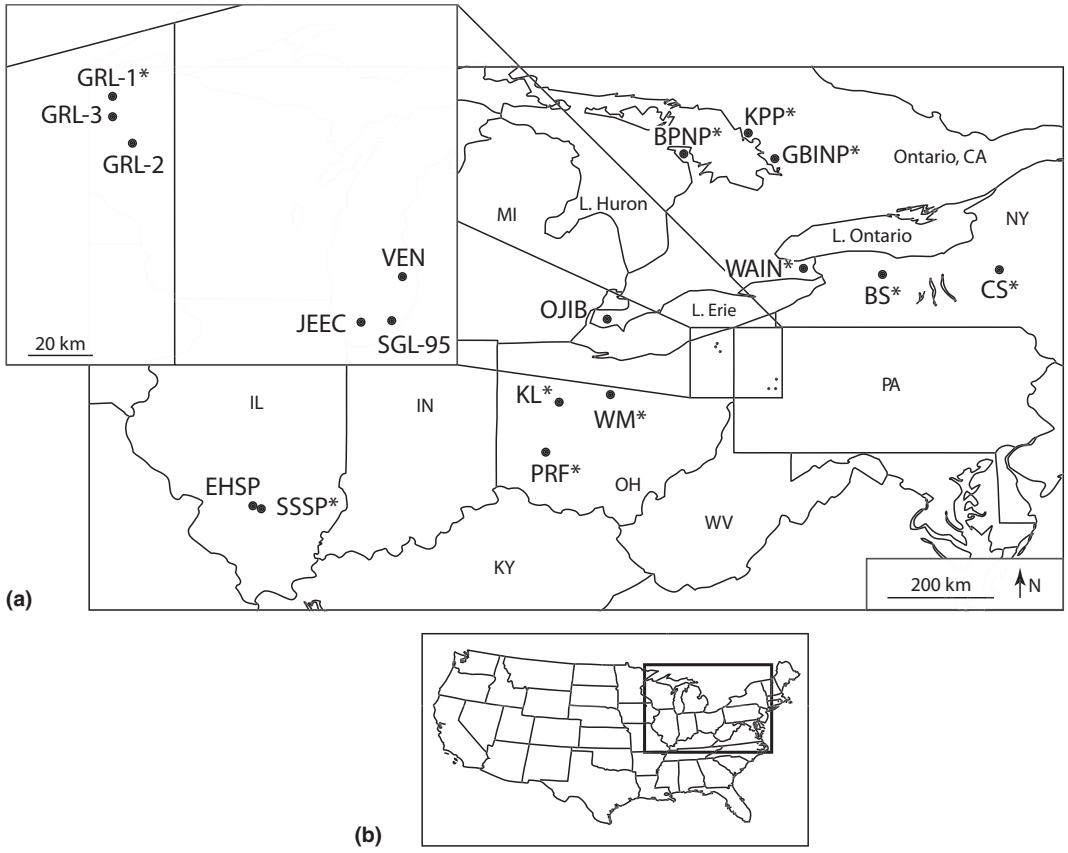

Fig. 1 (a) Sampling locations for eastern massasauga rattlesnake populations showing the Great Lakes and relevant state abbreviations. Population codes for the sampling locations are found in Table 1. Populations with asterisks are those used in the range-wide analysis using STRUCTURE. KL represents 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.

(Wilson & Rannala 2003) and MIGRATE (Beerli 2008), 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 timescale: 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 we 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 θ_{Ne} ($4N_e\mu$) and asymmetric gene flow M (m/μ) between pairs of populations over much longer (approximately $4N_e$ generations in the past; Beerli 2008) periods of time (approximately 1000's of years). As such, these migration rates provide estimates of gene flow that post- and pre-date the estimated time (approximately 300 yr BP: Mönkkönen & Welsh 1994) when humans began to significantly alter the habitats in which these snakes currently live.

For BAYESASS analyses, we 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 3×10^6 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), we 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, we then ran the analysis again using the seed from the best fit run but increased the run length to 3×10^7 iterations. The results presented are from this final run.

We used MIGRATE to estimate long-term estimates of gene flow and effective population size. As with the STRUCTURE analyses, we conducted the MIGRATE

analyses at both regional and range-wide spatial scales. At the regional scale, we estimated gene flow between the clusters of discrete sample locations all within 25 km of each other (see above). At the range-wide scale, we did not estimate migration rates for every possible population combination (Palstra *et al.* 2007). Rather, because of the limited gene flow suggested by high levels of differentiation (see below and Gibbs *et al.* 1997), we analysed 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 Georgian Bay Islands National Park (GBINP) because of 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, we 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 three 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 three 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 θ_{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 we assumed that the final run had converged and we present results from this final run. For all analyses, only 20 randomly chosen individuals were analysed per population because of computational demands and evidence that >20 individuals does not increase the accuracy of parameter estimates in MIGRATE analyses (Kuhner 2006).

To compare migration rates generated by BAYESASS and MIGRATE, we used the values of *m* directly generated by BAYESASS and estimated *m* from values of *M* (m/μ) generated by MIGRATE by dividing all *M* values by an estimated mutation rate of 5×10^{-4} (Garza & Williamson 2001). For this comparison, we excluded all estimates of migration rates associated with OJIB because of the inflated contemporary migration rates linked to this population (see below). To statistically test for similarity in migration rates, we 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

We used two methods implemented in BOTTLENECK to detect genetic bottlenecks. First, we 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, we 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 nonbottlenecked 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). We 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 timescale and were assessed for significance at the 0.05 level.

Tests of population history

Finally, we 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. We 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. We used Bayes factor to describe the probability of the most likely model over the probability of the other model. We 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

We genotyped 388 eastern massasaugas from 19 localities at 19 polymorphic microsatellite loci (Table 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 (Appendix S1, Supporting information).

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) (Appendix S1, Supporting information). 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 (six 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 ≤8% for all loci.

Population differentiation

Overall, 21% of the genetic variation present in the data set was because of among locality variation (θ_{Fst} = 0.206, 95% CI: 0.18-0.23, $P < 0.0001$). Pairwise θ_{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), θ_{Fst} was significant after applying the B–Y correction for multiple tests (Appendix S2, Supporting information), 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 which closely corresponded to the sampling localities with the exception of the two localities with the smallest sample sizes ($n \leq 9$) (OJIB and Jennings Environmental Education Center (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 above-mentioned, small sample size has been shown to reduce effectiveness of STRUCTURE in assigning individuals (Evanno *et al.* 2005) and so we assumed that the low assignment probabilities for these individuals were an artefact of low sample size and removed these populations from the analysis. With the modified data set containing 11 sampling localities, the Evanno method, determined the most likely value of *K* was 11 (K_{MAX} = 25.15) (Appendix S3, Supporting information). Of the five runs for $K = 11$, the run with the highest posterior probability ($\ln \Pr(X \mid K) = -14834.2$) had a highly nonrandom association of individuals corresponding strongly to sampling localities (Fig. 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 favoured, 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 $K = 2$ (ΔK_{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 $K = 3$ (ΔK_{MAX} = 441.4) and the three clusters were again defined by sampling locality. Finally, for the three populations in western Pennsylvania, a $K = 3$ received the greatest support (ΔK_{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,

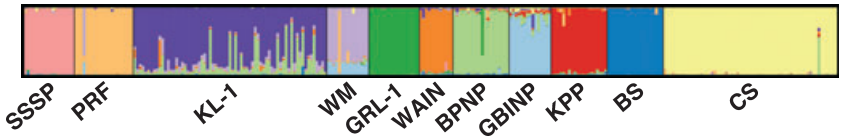

Fig. 2 Bar plot for STRUCTURE at K = 11 (ln P (D) = −14834) showing the association among genetic clusters and sample locations. Abbreviations for sampling locations are as follows: SSSP (South Shore State Park); PRF (Prairie Road Fen); KL-1 (Killdeer Plains Wildlife Area); WM (Willard Marsh Wildlife 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).

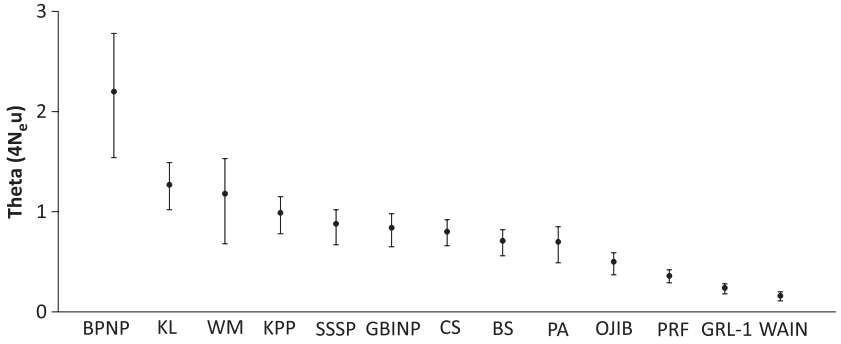

Fig. 3 Population estimates of theta (θ_{Ne}) from MIGRATE. Values are shown in decreasing order from left to right. Error bars represent the 95% confidence intervals for each estimate. KL represents KL-1.

KL-2 and WM) while PRF was among the least diverse populations in the study (H_E = 0.56). Estimates of θ_{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 (Fig. 3).

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 probably reflects real migration because of the close proximity of these two sites and the small nonsignificant θ_{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 in the following paragraphs, these results were interpreted as because of analytical artefacts.

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 to 11 (Fig. 4a). Of the 27 pairwise estimates of *M* for sets of populations, 10 were significantly asymmetrical (nonoverlapping 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_em (product of theta and *M* divided by four), ranged from 0.01 to 1.21. Our migration matrix resulted in 54 migration parameters among 13 populations and only four translated into N_em estimates >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 (Fig. 4b). Estimates of θ_{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_em was less than one (range: 0.64–0.79). θ_{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_em were <1 (range: 0.05–0.23). Similarly, in northeast Ohio, there was a 4-fold difference in θ_{Ne} among populations

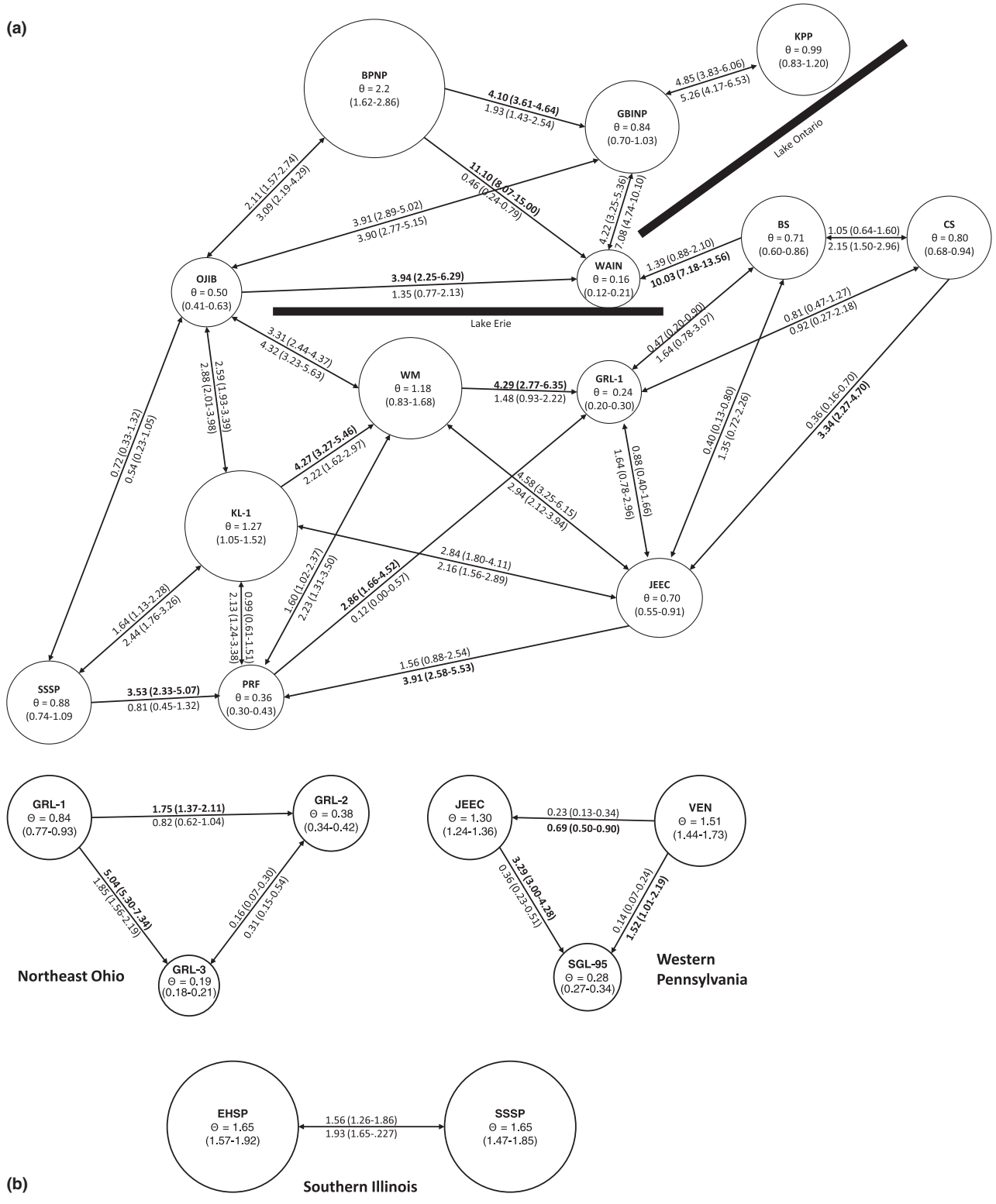

Fig. 4 Likelihood estimates of long-term gene flow among populations at (a) range-wide and (b) regional scales. Circle sizes reflect relative population size. Values inside the circles are θ_{Ne} shown with 95% confidence intervals. Bold values indicate migration was asymmetrical as determined from nonoverlapping 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.

(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_em from northeast Ohio were also <1 (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 (Fig. 5). 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 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). Under the TPM, only PRF showed evidence of a bottleneck (PRF, $P < 0.05$; Table 2). The mode-shift test deviated from the expected L-shaped distribution in 3 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 (approximately 100–200 years).

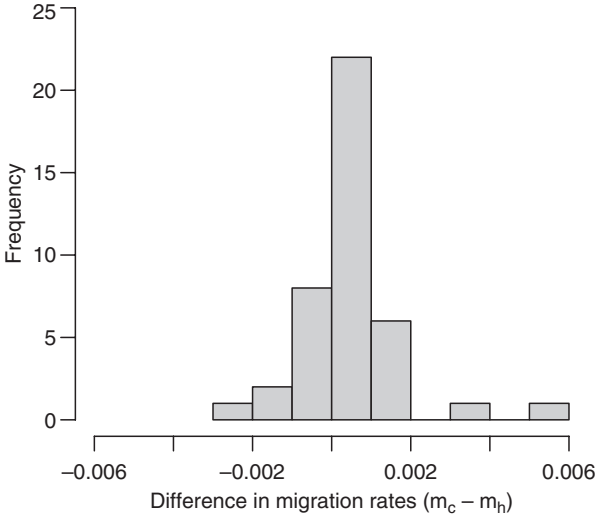

Fig. 5 Distribution of differences between contemporary migration rates (m_c) estimated using BAYESASS and historical migration rates (m_h) estimated using MIGRATE for specific population comparisons (m_c–m_h) as shown in Fig. 4.

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 1). Estimates of θ_{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

Our major results are that (i) Both contemporary and historical migration rates among populations were low

Table 2 Summary of results from BOTTLENECK for 16 populations for the Wilcoxon's test for both the two-phase model (TPM) and stepwise mutation 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

Population ID	Wilcoxon's test TPM	Wilcoxon's test SMM	Mode-shift test
SSSP	0.45	0.74	L-shape
EHSP	0.81	0.67	L-shape
PRF	0.00067	0.00019	L-shape
KL	0.61	0.34	L-shape
WM	0.76	0.46	L-shape
GRL-1	0.18	0.34	Shifted mode
GRL-2	0.71	0.55	L-shape
GRL-3	0.60	0.44	L-shape
JEEC	0.74	0.89	Shifted mode
BS	0.82	0.70	L-Shape
CS	0.34	0.63	L-shape
OJIB	0.71	0.82	Shifted mode
WAIN	0.15	0.07	L-shape
BPNP	0.57	0.14	L-shape
GBINP	0.15	0.04	L-shape
KPP	0.11	0.04	L-shape

and similar in magnitude even for populations located less than 7 km apart, and a test of alternate models of population history strongly favours a model of long-term drift-migration equilibrium over a recent isolation drift-only model; (ii) Geographically discrete massasauga populations exhibit a high degree of genetic structure even at regional spatial scales; and (iii) There is a twofold difference in genetic effective population sizes but little genetic evidence that populations have undergone recent or historical declines in size. We discuss the evolutionary and conservation implications of these findings below.

Contemporary and historical gene flow

Our 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 because 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 pre-date colonization and subsequent landscape modification by European settlers in North America (Schmidt 1938; Pielou 1991). 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 (Reinert & Kodrich 1982; Johnson 2000). 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. Our results add to the increasing number of studies that emphasize the importance of historical factors in moulding patterns of genetic variation within and between populations of species that occupy highly impacted landscapes (Cunningham & Moritz 1998; Vandergast *et al.* 2007; Hansen *et al.* 2009; Pavlacky *et al.* 2009).

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 because of the repeated exposure of deleterious recessive alleles to selection; although behavioural 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 (Chiucchi & Gibbs, unpublished data). 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 raise the question of whether these results could be an artefact 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 θ and thereby affect migration estimates (Beerli 2004). However, we 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). 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 our 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 (Prior

et al. 1997; Bushar *et al.* 1998; Prosser *et al.* 1999; Manier & Arnold 2005; Paquin *et al.* 2006; Clark *et al.* 2008; Marshall *et al.* 2009; 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). Our 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

Our estimates of genetic effective population size from the MIGRATE analyses (θ_{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 because of 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 θ_{Ne} (all >1.0) hence the highest levels of standing neutral variation among those analysed. 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 θ_{Ne} (e.g. PRF, GRL-1 and WAIN) may now be recognized as potentially being at increased risk of extinction because of loss of adaptive genetic variation.

We caution that use of θ_{Ne} to estimate levels of standing adaptive variation in eastern massasauga populations relies on a number of important assumptions. First, θ_{Ne} is assumed to be based on standing genetic variation from a single interbreeding deme of snakes. We 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 θ_{Ne} for populations showing high values for this parameter (e.g. BPNP). Second, and most significantly, we 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 θ_{Ne} show no significant correlation with population level variation in an adaptive trait (venom variation) which is at least partially under genetic control (Gibbs & Chiucchi, submitted; 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, unpublished data). 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 2003).

A second use for population-specific values of θ_{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 because of 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 (Seigel & Sheil 1999; Middleton & Chu 2004; Miller 2005). 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 <0.5 (Frankham 1995) with the best estimates for vertebrates averaging 0.38 (Vucetich *et al.* 1995). If we assume a widely used mutation rate of 5×10^{-4} mutations/generation/locus (Garza & Williamson 2001) to convert estimates of θ_{Ne} to N_e, then estimates of N_e range from approximately 80 (WAIN) to 1100 (BPNP) individuals. Assuming a ratio of $N_e/N_c = 0.38$, then we estimate that N_c for populations surveyed here range from approximately 200 to 2900 snakes. If we use a more conservative ratio of $N_e/N_c = 0.5$, then all populations have N_c values of ≥160 individuals. Although estimated values of N_c are crucially dependent on the mutation rate and ratio of N_e/N_c used, our results suggest that under a variety of parameter values census sizes of all populations are >160 individuals. This is significant, as our estimate is

greater than the minimum population size (approximately 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 (Seigel & Sheil 1999; Middleton & Chu 2004). We caution, however, that a crucial assumption with this use of genetically based measures of θ_{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.

Acknowledgements

This study would not have been possible without the very generous assistance with the collection of samples by many people over a number of years. Many individuals have helped but in particular, we thank Jeff Davis, Michael Dreslik, Brian Fedorko, Tony Frazier, Kim Frolich, Dan Harvey, Matt Kowalski, Greg Lipps, Chris Parent, Paul Pratt, Kent Prior, Kevin Shoemaker, Michelle Villenueve, Pat Weatherhead and Doug Wynn for their enthusiastic help with finding and collecting samples from rattlesnakes. We also thank Jose Diaz, Christine Anderson and Mike Murphy for assistance with the laboratory work, Peter Beerli and Laura Kubatko for advice on analyses, Katy Greenwald for comments on the paper and Carolyn Caldwell (Wildlife Division, Ohio Department of Natural Resources) for her long-term, sustained support of this work. This research was supported by funds from the Wildlife Division, Ohio Department of Natural Resources and Ohio State University.

References

Anderson CD, Gibbs HL, Douglas ME, Holycross AT (2009) Conservation genetics of the desert massasauga rattlesnake (*Sistrurus catenatus edwardsii*). *Copeia*, **2009**, 740–747.

Anderson CS, Gibbs HL, Chiucchi J (2010) Nineteen polymorphic microsatellite loci isolated from the Eastern Massasauga Rattlesnake, *Sistrurus c. catenatus*. *Conservation Genetics* DOI: 10.1007/s12686-010-9253-y.

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, **13**, 827–836.

Beerli P (2008) Migrate version 3.0: a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Distributed over the internet at http://popgen.scs.edu/migrate.html.

Beerli P (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use. In:*Population Genetics for Animal Conservatio* (Eds Bertorell G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C). pp. 42–79, Cambridge University Press, Cambridge.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA*, **98**, 4563–4568.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, **29**, 1165–1188.

Bonnet E, Van de Peer Y (2002) zt: a software tool for simple and partial Mantel tests. *Journal of Statistical Software*, **7**, 1–12.

Bushar LM, Reinert HK, Gelbert L (1998) Genetic variation and gene flow within and between local populations of the timber rattlesnake, *Crotalus horridus*. *Copeia*, **2**, 411–422.

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

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–631.

Ciofi C, Beaumont MA, Swingland IR, Bruford MW (1999) Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. *Proceedings of the Royal Society of London Series B*, **266**, 2269–2274.

Clark RW (2004) Kin recognition in rattlesnakes. *Proceedings of the Royal Society of London Series B Biology Letters*, **271**, S243–S245.

Clark RW, Brown WS, Stechert R, Zamudio KR (2008) Integrating individual behaviour and landscape genetics: the population structure of timber rattlesnake hibernacula. *Molecular Ecology*, **17**, 719–730.

Cook FR (1992) After an ice age: zoogeography of the massasauga within a Canadian herpetofaunal perspective. In:*International Symposium and Workshop on the Conservation of the Eastern Massasauga Rattlesnak* (eds Johnson B, Menzies V). pp. 19–25, Zoo, Toronto, ON.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cunningham M, Moritz C (1998) Genetic effects of forest fragmentation on a rainforest restricted lizard (Scincidae: *Gnypetoscincus queenslandiae*). *Biological Conservation*, **83**, 19–30.

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Earl D (2009) Structure Harvester v0.3, from website: http://users.soe.ucsc.edu/~dearl/software/struct_harvest/

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

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

Frankham R (1995) Conservation Genetics. *Annual Review of Genetics*, **29**, 305–327.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*, Cambridge University Press, Cambridge.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.

Gibbs HL, Rossiter W (2008) Rapid evolution by positive selection and gene gain and loss: PLA_2 venom genes in closely related Sistrurus rattlesnakes with divergent diets. *Journal of Molecular Evolution*, **66**, 151–166.

Gibbs HL, Weatherhead PJ (2001) Insights into population ecology and sexual selection in snakes through the application of DNA-based genetic markers. *Journal of Heredity*, **92**, 173–179.

Gibbs HL, Prior KA, Weatherhead PJ, Johnson G (1997) Genetic structure of populations of the threatened eastern massasauga rattlesnake, *Sistrurus c. catenatus*: evidence from microsatellite DNA markers. *Molecular Ecology*, **6**, 1123–1132.

Gibbs HL, Sanz L, Calvete J (2009) Snake population venomics: Proteomics-based analyses of individual variation reveals gene regulation effects on venom protein expression in *Sistrurus* rattlesnakes. *Journal of Molecular Evolution*, **68**, 113–125.

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.

Government of Canada (2009) Species at Risk Public Registry. http://www.sararegistry.gc.ca.

Hansen BD, Harley KP, Lindenmayer DB *et al.* (2009) Population genetic analysis reveals a long-term decline of a threatened endemic Australian marsupial. *Molecular Ecology*, **18**, 3346–3362.

Hartl GB, Zachos FE, Nadlinger K *et al.* (2005) Allozyme and mitochondrial DNA analysis of French red deer (*Cervus elaphus*) populations: genetic structure and its implications for management and conservation. *Mammalian Biology*, **70**, 24–34.

Johnson G (2000) Spatial ecology of the Eastern Massasauga *Sistrurus c. catenatus*) in a New York peatland. *Journal of Herpetology*, **34**, 186–192.

Johnson JA, Toepfer JE, Dunn PO (2003) Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie chickens. *Molecular Ecology*, **12**, 3335–3347.

Keyghobadi N (2007) The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, **85**, 1049–1064.

Kubatko LS, Gibbs HL, Bloomquist EW (2010) Inferring species-level phylogenies and taxonomic distinctiveness using multi-locus data in *Sistrurus* rattlesnakes. *Systematic Biology* (in press).

Kuhner M (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics*, **22**, 768–770.

Loader C (1999) *Local Regression and Likelihood*, Springer Mathematics, New York.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238–247.

Manier MK, Arnold SJ (2005) Population genetic analysis identifies source-sink dynamics for two sympatric garter snake species (*Thamnophis elegans* and *Thamnophis sirtalis*). *Molecular Ecology*, **14**, 3965–3976.

Marshall JCJ, Kingsbury BA, Minchella DJ (2009) Microsatellite variation, population structure, and bottlenecks in the threatened copperbelly water snake. *Conservation Genetics*, **10**, 465–476.

Martinez-Cruz B, Godoy JA, Negro JJ (2007) Population fragmentation leads to spatial and temporal genetic structure in the endangered Spanish imperial eagle. *Molecular Ecology*, **16**, 477–486.

Martinez-Solano I, Rey I, Garcia-Paris M (2005) The impact of historical and recent factors on genetic variability in a mountain frog: the case of *Rana iberica* (Anura: Ranidae). *Animal Conservation*, **8**, 431–441.

Middleton J, Chu JY (2004) Population viability analysis (PVA) of the Eastern Massasauga Rattlesnake, *Sistrurus catenatus catenatus*, in Georgian Bay Islands National Park and elsewhere in Canada. Prepared for the Eastern Massasauga Rattlesnake Species Recovery Team, January 2004. 52pp.

Miller P (2005) Population viability assessment for the Eastern Massasauga Rattlesnake (*Sistrurus catenatus catenatus*) on the Bruce Peninsula, Ontario, Canada. In: *Third International Symposium and Workshop on Conservation of the Massasauga (Sistrurus catenatus): Population Viability and Outreach* (eds Massasauga Recovery Team). 2005 October 12-14, pp. 37. Toronto Zoo, Toronto, ON.

Mönkkönen M, Welsh DA (1994) A biogeographical hypothesis on the effects of human caused landscape changes on the forest bird communities of Europe and North America. *Annales Zoologici Fennici*, **31**, 61–70.

Narum SR (2006) Beyond Bonferroni: less conservative analyses for conservation genetics. *Conservation Genetics*, **7**, 783–787.

Palstra FP, O'Connell MF, Ruzzante DE (2007) Population structure and gene flow reversals in Atlantic salmon (*Salmo trutto*) over contemporary and long-term temporal scales: effects of population size and life history. *Molecular Ecology*, **16**, 4504–4522.

Paquin MM, Wylie GD, Routman EJ (2006) Population structure of the giant garter snake, *Thamnophis gigas*. *Conservation Genetics*, **7**, 25–36.

Pavlacky Jr DC, Goldizen AW, Prentis PJ *et al.* (2009) A landscape genetics approach for quantifying the relative influence of historic and contemporary habitat heterogeneity on the genetic connectivity of a rainforest bird. *Molecular Ecology*, **18**, 2945–2960.

Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: present problems, paradoxes and future challenges. *Biodiversity and Conservation*, **16**, 4147–4163.

Pielou EC (1991) *After the Ice Age: The Return of Life to Glaciated North America*, University of Chicago Press, Chicago, Illinois.

Piry S, Luikart G, Cornuet JM (1999) Computer note. Bottleneck: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Prior KA, Gibbs HL, Weatherhead PJ (1997) Population genetic structure in the black rat snake: implications for management. *Conservation Biology*, **11**, 1147–1158.

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

Prosser MR, Gibbs HL, Weatherhead PJ (1999) Microgeographic population genetic structure in the

northern water snake, *Nerodia sipedon sipedon* detected using microsatellite DNA loci. *Molecular Ecology*, **8**, 329–333.
Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
Redeker S, Andersen LW, Pertoldi C *et al.* (2006) Genetic structure, habitat fragmentation and bottlenecks in Danish bank voles (*Clethrionomys glareolus*). *Mammalian Biology*, **71**, 144–158.
Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. *Conservation Biology*, **17**, 230–237.
Reinert HK, Kodrich WR (1982) Movements and habitat utilization by the massasauga, *Sistrurus catenatus catenatus*. *Journal of Herpetology*, **16**, 162–171.
Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
Schmidt KP (1938) Herpetological evidence for the post-glacial extension of the steppe in North America. *Ecology*, **19**, 396–407.
Schwartz TS, Karl SA (2005) Population and conservation genetics of the gopher tortoise (*Gopherus polyphemus*). *Conservation Genetics*, **6**, 917–928.
Seigel RA, Sheil CA (1999) Population viability analysis: applications for the conservation of massasaugas. In:''*Second International Symposium and Workshop on the Conservation of the Eastern Massasauga Rattlesnake, Sistrurus Catenatus Catenatus: Population and Habitat Management Issues in Urban, Bog, Prairie and Forested Ecosystems* (eds Johnson B, Wright M). pp. 17–22, Toronto Zoo, Toronto, Ontario.
Spiegelhalter DJ (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B (Methodological)*, **64**, 583–639.
Szymanski J (1998) Status Assessment for Eastern Massasauga. U.S. Fish and Wildlife Service Series Status Assessment for Eastern Massasauga.
Ursenbacher S, Monney JC, Fumagalli L (2009) Limited genetic diversity and high differentiation among the remnant adder (*Vipera berus*) populations in the Swiss and French Jura Mountains. *Conservation Genetics*, **10**, 303–315.
Vandergast AG, Bohonak AJ, Weissman DB *et al.* (2007) Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: Stenpelmatus). *Molecular Ecology*, **16**, 977–992.
Vucetich JA, Waite TA, Nunney L (1995) Fluctuating population size and the ratio of effective to census population size. *Evolution*, **51**, 2017–2021.
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, **163**, 1177–1191.

James Chiucchi is a graduate student interested in using molecular methods to answer questions about the evolution and conservation of reptiles and amphibians. Lisle Gibbs uses molecular genetic techniques to study a broad range of questions about the evolution, ecology and conservation of snakes, salamanders and birds.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Summary of 19 polymorphic microsatellite loci. Null alleles were estimated using FreeNa.

Appendix S2 Genetic estimates of pairwise FST between pairs of populations calculated using MSA based on 10 000 permutations.

Appendix S3 ΔKcalculated with the Evanno method (2005) from the range-wide STRUCTURE analysis.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.