![USGS logo] **USGS**
science for a changing world

Prepared for the U.S. Fish and Wildlife Service

Conservation Genomics of the Mogollon Narrow-Headed Gartersnake (*Thamnophis rufipunctatus*) and Northern Mexican Gartersnake (*Thamnophis eques megalops*)

Open-File Report 2018–1141

U.S. Department of the Interior
U.S. Geological Survey

Conservation Genomics of the Mogollon Narrow-Headed Gartersnake (*Thamnophis rufipunctatus*) and Northern Mexican Gartersnake (*Thamnophis eques megalops*)

By Dustin A. Wood, Iain D. Emmons, Erika M. Nowak, Bruce L. Christman, Andrew T. Holycross, and Amy G. Vandergast

Prepared for the U.S. Fish and Wildlife Service

Open-File Report 2018 –1141

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia: 2018

Acknowledgments

Contents

Figures

Tables

Conversion Factors

International System of Units to Inch/Pound

Multiply	By	To obtain
Mass		
nanogram (ng)	3.527×10^{-11}	ounce, avoirdupois (oz)

Abbreviations

bp	base pairs
BIC	Bayesian Information Criterion
ddRAD	double-digest restriction site associated DNA
DAPC	Discriminant Analysis of Principal Components
DNA	deoxyribonucleic acid
HWE	Hardy-Weinberg equilibrium
MCMC	Markov chain Monte Carlo
ML	maximum likelihood
NGS	next-generation sequencing
PCR	polymerase chain reaction
RADseq	Restriction site associated DNA sequencing
SNP	single nucleotide polymorphism
US	United States
USFWS	US Fish and Wildlife Service
USGS	US Geological Survey
VCF	variable call format

This page left intentionally blank

Conservation Genomics of the Mogollon Narrow-Headed Gartersnake (*Thamnophis rufipunctatus*) and Northern Mexican Gartersnake (*Thamnophis eques megalops*)

By Dustin A. Wood, Iain D. Emmons, Erika M. Nowak, Bruce L. Christman, Andrew T. Holycross, and Amy G. Vandergast

Abstract

The ability of populations to persist and adapt to abiotic and biotic changes is reliant on genetic diversity. When connectivity across a species landscape is disrupted, the levels and distribution of genetic diversity can rapidly deteriorate as a result of genetic drift, leading to increased inbreeding and reduced adaptive potential. Therefore, understanding the distribution and degree of genetic variation within imperiled populations provides important information for conservation management and recovery strategies, especially when paired with translocation and repatriation programs. Here, we used genome-wide nuclear markers to study the population structure and genetic diversity from tissue samples collected between 2010 and 2016 of two threatened species of gartersnakes inhabiting the lower Colorado River Basin in the United States: Mogollon Narrow-headed gartersnake (*Thamnophis rufipunctatus*) and Northern Mexican gartersnake (*Thamnophis eques megalops*). Our specific objectives were to determine how populations inhabiting the lower Colorado River Basin were related to sister species and southern populations along the Sierra Madre Occidental in Mexico, to determine how genetic variation is partitioned among drainage basins in the lower Colorado River Basin, and to provide estimates of genetic diversity and effective sizes of sampled sites to aide species-specific conservation management of these threatened gartersnakes. For both species, we found that populations along the lower Colorado River Basin are highly differentiated from sister species and southern populations located further south in Mexico, and exhibit reduced genetic diversity relative to populations along the Sierra Madre Occidental. Within the lower Colorado River Basin, genetic analyses revealed highly structured genetic groups for both species of gartersnakes that point to shared contemporary and historical drivers of differentiation. We found that most sites throughout the lower Colorado River Basin have low genetic diversity and effective population sizes below the threshold required to retain adaptive potential. However, these trends were especially pronounced for *T. rufipunctatus*. If genetic management and translocation strategies are adopted in the future, these population genetic results can be used to highlight sites of particular concern and locate the most genetically similar sites for translocation or re-establishment efforts. Such measures could help curb further population genetic change, alleviate problems associated with low genetic diversity, and strengthen the adaptive potential across the range of these two gartersnake species.

Introduction

The Mogollon Narrow-headed gartersnake (*Thamnophis rufipunctatus*) and the Northern Mexican gartersnake (*Thamnophis eques megalops*) are highly aquatic species inhabiting the major perennial drainages throughout the lower Colorado River Basin in central and southern Arizona, western New Mexico, and for *T. e. megalops*, Mexico (Rossman and others, 1996; Wood and others, 2011). Although historically common throughout their range, recent surveys in the United States indicate that the amount of occupied habitat for both gartersnake species has declined greatly. The U.S. Fish and Wildlife Service (USFWS) estimates that 76 percent of *T. rufipunctatus* populations and 83 percent of *T. e. megalops* populations are likely not viable and may exist at low population densities that could be threatened with extirpation or may already be extirpated (U.S. Fish and Wildlife Service, 2013, 2014). Predation and displacement by non-native species appear to be the primary driving factors behind the range-wide declines, although they may be less severe for *T. eques megalops* (Emmons and others, 2016). Other potential drivers, such as water diversion, increased river sedimentation, severe floods, and catastrophic fires (Jennings and Christman, 2015; Nowak and others, 2017) have contributed to the loss and degradation of habitat that has disrupted drainage basin connectivity in the region (Rosen and Schwalbe, 1988; Rosen, 1993; Holycross and others, 2006). The cumulative effect of these threats has led the USFWS to list both species as threatened, range-wide, under the Endangered Species Act, and to propose critical habitat (U.S. Fish and Wildlife Service, 2013, 2014).

Despite this protection, little is known about the genetic diversity in existing populations, information that is useful in designing recovery plans to ensure the long-term viability of these snakes (Elgar and Clode, 2001). Extensive threats across the ranges of these gartersnakes have likely lead to reduced genetic exchange and small population sizes, which contribute to loss of genetic diversity and inbreeding, and can reduce the ability to adapt to environmental change (Shaffer, 1981; Frankham, 2005). At present, too little is known about the population genetics of either gartersnake species to make informed conservation management decisions that would further protect and maintain maximum genetic diversity and historical population structure. Recent phylogenetic work involving both species has examined species-level relationships among other North American gartersnakes (Alfaro and Arnold, 2001; de Quieroz and others, 2002; Wood and others, 2011; McVay and others, 2015), but none of the work has explicitly examined the range-wide population genetic structure and diversity. Information on the distribution and degree of genetic variation within populations can also inform repatriation and genetic rescue efforts (Frankham and others, 2017). When integrated with other species-specific management actions aimed at removing threats and improving demographic viability, genetic rescue and repatriation could help curb further population genetic change, alleviate problems associated with low genetic diversity, and strengthen adaptive potential across the range of these two gartersnake species.

In this study, we used a restriction-site associated next-generation sequencing (RADseq) approach to address three objectives. Our first objective was to compare the lower Colorado River Basin distributions of *T. rufipunctatus* and *T. e. megalops* with their counterparts that range farther south in Mexico to evaluate the divergence and genetic diversity differences between the lower Colorado River Basin populations and those along the Sierra Madre Occidental. Previous phylogenetic work (Wood and others, 2011) revealed *T. rufipunctatus* to be a distinct lineage from sister populations along the Sierra Madre Occidental supporting full species status, but lineage divergence has never been investigated using genetic data across the range of *T. e. megalops*. Our second objective was to quantify spatial patterns of genetic structure in the lower Colorado River Basin for each species to evaluate whether drainage basins are consistent with major genetic groups. Connectivity between the Salt and Gila Rivers and their associated drainage basins has not been previously investigated but is critical to understand for conservation planning. Our last objective was to estimate genetic diversity indices and effective population sizes for sites sampled across the two species' ranges and evaluate them against established thresholds (Frankham and others, 2014; Frankham and others, 2017) to identify locations that may be more susceptible to negative genetic factors such as inbreeding and genetic drift. The work here is intended to contribute to developing management plans for each species, specifically by identifying the genetic factors that could be used in selection of appropriate sources for existing and future captive breeding research programs as well as critical locations for source and recipient conservation translocations. These genetic metrics also provide a baseline for future monitoring of population status and the effectiveness of management actions.

Methods

Samples were collected for this study during the spring and summers between 2010 and 2016, in addition to samples provided by Arizona Department of Game and Fish, Phoenix Zoo, Albuquerque Biological Park, Museum of Southwestern Biology at University of New Mexico, and deoxyribonucleic acid (DNA) samples from a previous study (Wood and others, 2011). In total, we acquired 351 gartersnake tissue samples (tail-tips and (or) blood) from most occupied drainages throughout Arizona and New Mexico, although we did not sample potentially occupied drainages within Tribal lands (White Mountain Apache, San Carlos Apache, and Tohono O'odham Reservations). The samples included 192 *T. rufipunctatus* from 11 sites throughout the Verde River, Salt River, San Francisco River, and Gila River Basins (fig. 1), and 159 *T. e. megalops* from 10 sites throughout the Bill Williams River, Verde River, Salt River, Santa Cruz River, and Gila River Basins (fig. 2). However, different subsets of samples were used in different analyses (details of which are listed under each objective and in the appendixes). For *T. rufipunctatus*, we restricted genetic diversity and population structure analyses to extant sites with sample size ≥ 3, so Whitewater and Tonto Creek were removed from these analyses. We also included samples from sister species (for *T. rufipunctatus*) or sister populations (for *T. e. megalops*) that were present farther south along the Sierra Madre Occidental in Mexico to test for phylogenetic divergence and compare genetic variation levels with United States populations (fig. 3).
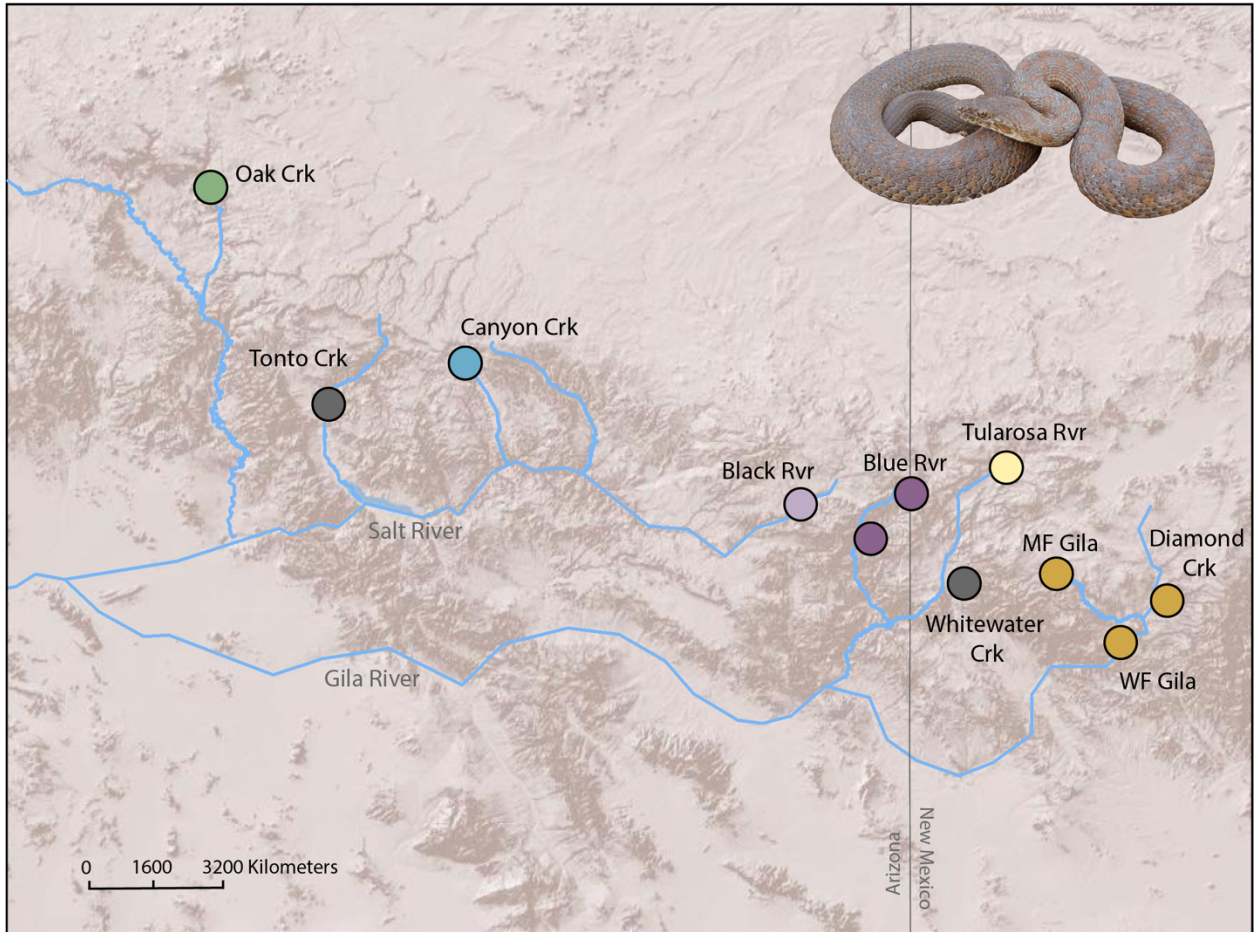
Figure 1. Sampling map of *Thamnophis rufipunctatus* used in this study across the lower Colorado River Basin. Sites are colored according to the genetic clusters identified in the study, both Tonto Creek and Whitewater Creek were not included in population structure and genetic diversity analyses because of low sample sizes and (or) insufficient data. Abbreviations used: Crk = Creek, Rvr = River, MF = Middle Fork, WF = West Fork.
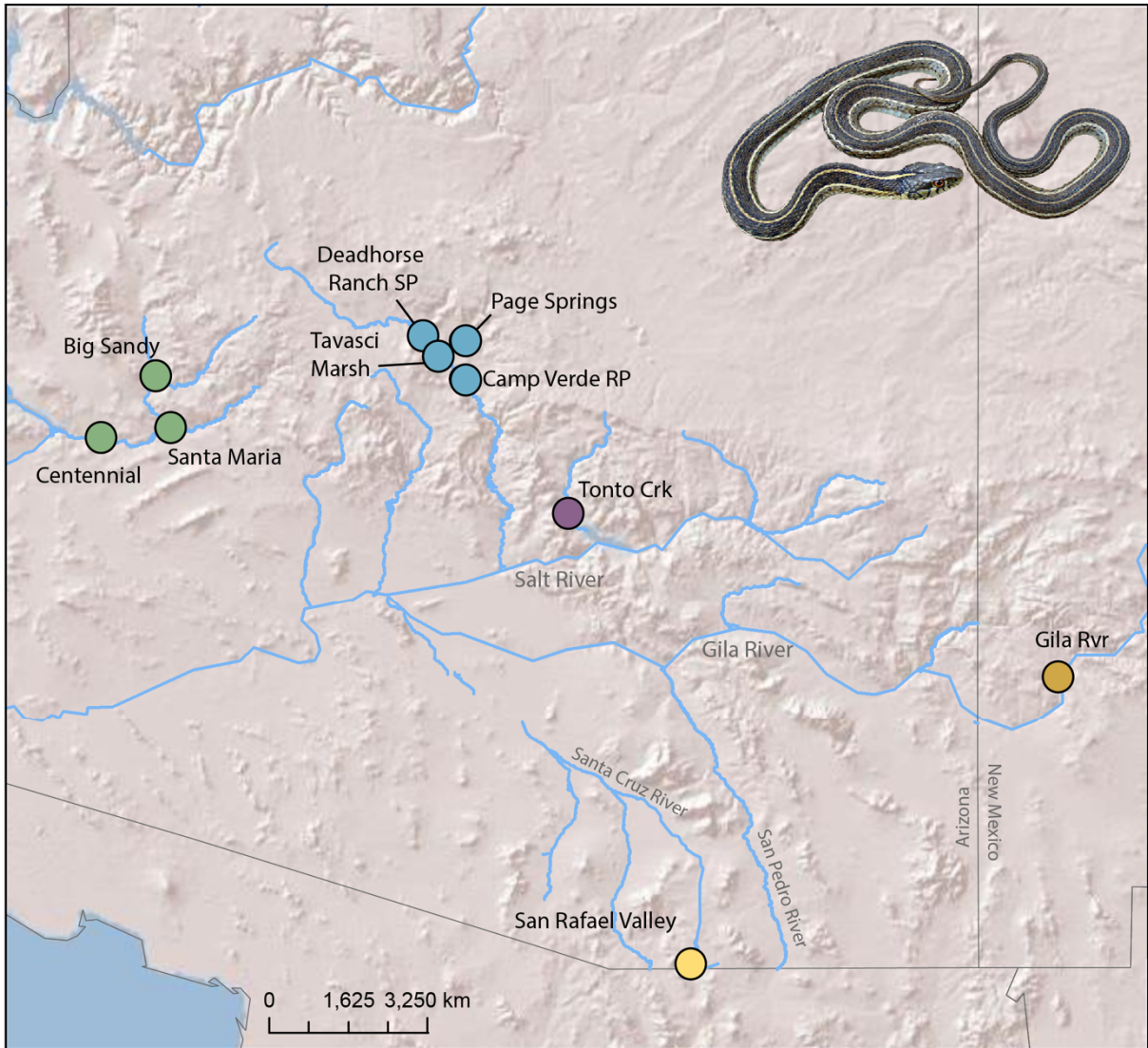
Figure 2. Sampling map of *Thamnophis eques megalops* used in this study across the lower Colorado River Basin. Sites are colored according to the genetic clusters identified in the study. Abbreviations used: Crk = Creek, Rvr = River, RP = Regional Park, SP = State Park.
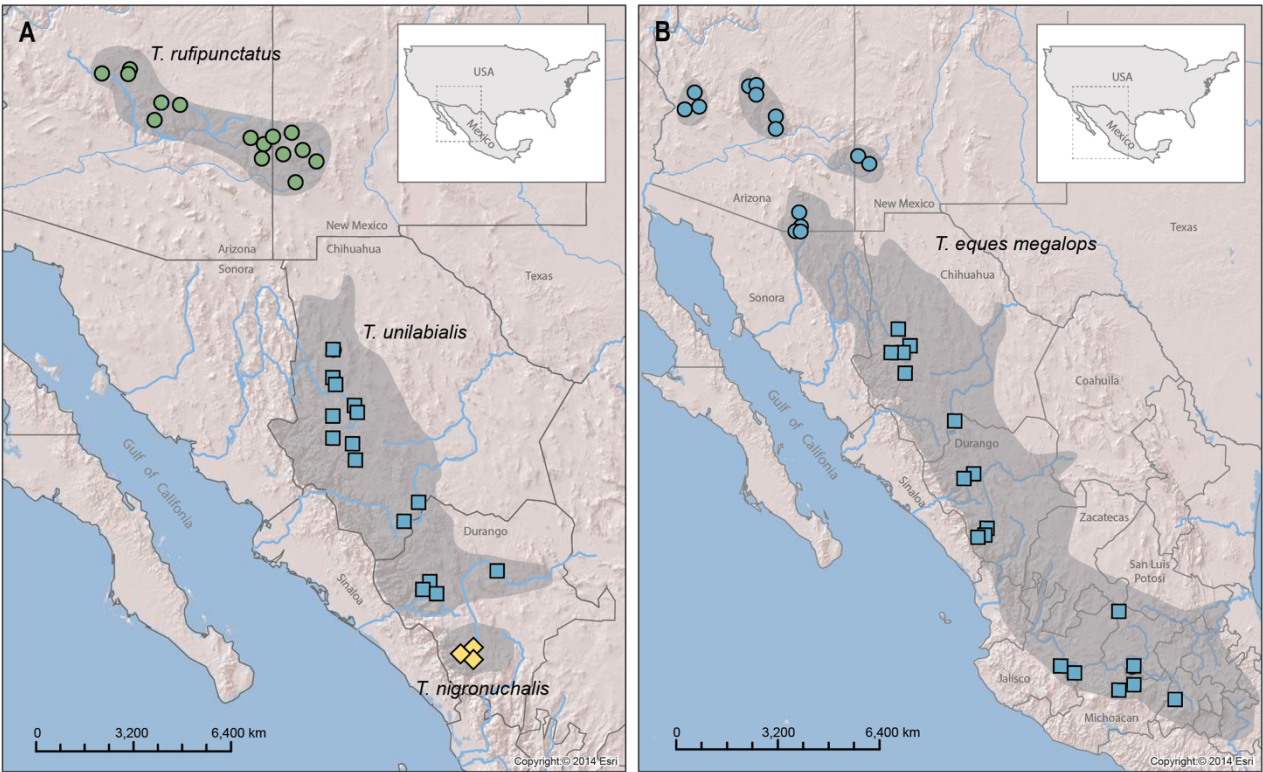
Figure 3. Range of each species (grey shadow) and distribution of samples used in this study for the (a) Narrow-headed gartersnake species complex: *Thamnophis rufipunctatus* (circles), *Thamnophis unilabialis* (squares), *Thamnophis nigronuchalis* (diamonds), and (b) the Northern Mexican gartersnake: *Thamnophis eques megalops*, samples within the lower Colorado River Basin (circles) and the Mexican highlands (squares).

Data Generation

All tissues were preserved in 95-percent ethanol after collection and genomic DNA was extracted using Qiagen DNeasy Blood and Tissue Kits (Qiagen, Valencia, California). Prior to next-generation sequencing (NGS) library preparation, we quantified DNA on a Qubit fluorometer (Life Technologies), and 500–700 nanograms (ng) of DNA were used for library preparation. We followed the double-digest restriction-associated DNA (ddRAD) sequencing protocol developed in Peterson and others (2012) for NGS library preparation, with some modifications. We digested genomic DNAs using 20 units each of the restriction enzymes *Sbf*I and *Msp*I (New England Biolabs, U.S.A.) and used Agencourt AMPure beads (Beckman Coulter, Danvers, Massachusetts) to purify the digestions prior to ligating uniquely bar-coded adapters with T4 ligase (New England Biolabs). We quantified all ligation products on the Qubit fluorometer, pooled across 12 index groups in equimolar concentrations, and then size selected fragments between 415 and 515 base pairs (bp) using a Pippin Prep size fractionator (Sage Science, Beverly, Mass.). We amplified the recovered fragments from each pool using 5–10 ng of the recovered DNA, Phusion High-Fidelty *Taq* (New England Biolabs), and Illumina's primers. Polymerase chain reaction (PCR) products were then cleaned with Agencourt AMpure beads (Beckman Coulter) and quantified using the Qubit fluorometer (Life Technologies) and Bioanalyzer (Agilent Technologies, U.S.A.) before being pooled for sequencing (50 bp single end reads) on an Illumina HiSeq 4000 at the Vincent J. Coates Genomic Sequencing Laboratory at University of California, Berkeley.

6

Bioinformatics

We used two bioinformatics pipelines to process the NGS libraries. We used the PYRAD pipeline (Eaton, 2014) for phylogenetic analyses because this pipeline can incorporate insertion-deletion polymorphisms when building loci and is generally used in phylogenetic applications. To demultiplex and process the raw sequence reads, we first used PYRAD to trim adapter sequences and enzyme cut sites, and then discarded any low-quality reads with Phred scores < 20. We required a minimum of 10 reads for a locus to be called within an individual. We clustered reads using the program vsearch (www.github.com/torognes/vsearch) and aligned the sequences using muscle (Edgar, 2004) with a clustering threshold of 90 percent. We discarded consensus sequences with more than 20 percent heterozygous sites and more than two alleles for an individual. For each species of gartersnake, we generated multiple datasets composed of different subsets of individuals and different thresholds of missing data depending on the objective being evaluated, and we report these differences below.

For population structure and genetic diversity objectives, we generated datasets using the STACKS version 1.46 (Catchen and others, 2013) bioinformatics pipeline. For each species, we moved the demultiplexed raw reads into STACKS and executed the ustacks, cstacks, and sstacks modules using the *denovo map* wrapper under the following parameters: minimum stack depth (m) of 3, mismatch distance between loci within an individual (M) of 2, and number of mismatches between loci in the catalogue (n) of 1 (Paris and others, 2017). All loci and variable sites produced by STACKS were subjected to a final filter approach that retained loci present across at least seven sampled sites and were found in 80 percent of the individuals sequenced with a minimum allele frequency cut off set at 5 percent.

Objective 1—Testing for Phylogenetic Divergence of United States Populations

To test for divergence of United States populations of gartersnakes relative to sister species (for *T. rufipunctatus*) or sister populations (for *T. eques megalops*) ranging farther south along the Sierra Madre Occidental, we conducted phylogenetic analyses on RADseq loci generated from PYRAD that included samples for Sierra Madre Occidental for each species. For tests with *T. rufipunctatus*, we included 42 samples of *T. rufipunctatus* from the major drainages within the lower Colorado River Basin, 22 samples of *T. unilabialis* from Chihuahua and Durango, Mexico, and 4 samples of *T. nigronuchalis* from Durango, Mexico. For tests involving *T. e. megalops*, we included 54 samples from the major drainages within the lower Colorado River Basin, 33 samples of *T. e. megalops* from Chihuahua and Durango, Mexico, and 16 samples of *T. eques* from Jalisco, Michoacan, and Guanajuato, Mexico. For phylogenetic analyses, we used the Bayesian Markov chain Monte Carlo (MCMC) approach in mrbayes version 3.2 (Ronquist and others, 2011) and beast 1.8.4 (Drummond and others, 2012). For each analysis, MCMC searches of tree space included four independent runs with one (beast) to four (mrbayes) Markov chains each using default heating values, and each were run for 10 million generations. Trees were sampled every 1,000th generation, resulting in 10,000 trees from each run. We assessed evidence for convergence using tracer 1.6 (Rambaut and others, 2014). We considered lineages with posterior probabilities of ≥ 0.95 to be strongly supported.

Objective 2—Quantify Spatial Patterns of Genetic Structure

For both species of gartersnakes, we evaluated population genetic structure with multiple analytical methods. First, we used SNAPCLUST (Beugin and others, 2018), then implemented in the R package ADEGENET (Jombart and others, 2010; Jombart and Ahmed, 2011), to estimate the most likely number of panmictic populations or genetic clusters given the data. SNAPCLUST uses a two-step maximum likelihood (ML) approach to compute ML estimates of clusters from the allele frequencies, assuming Hardy-Weinberg equilibrium (HWE). SNAPCLUST relies on goodness-of-fit statistics (Bayesian information criterion, BIC) to help guide selection of the optimal number of genetic clusters. To compliment the SNAPCLUST analyses, we used Discriminant Analysis of Principal Components (DAPC), a multivariate ordination approach implemented in the R package ADEGENET version 2.1.1. This method evaluates the optimal number of genetic clusters using PCA ordination to maximize the between-group variation while minimizing the variation found within groups and does not require the assumption of HWE or unlinked markers. We used information from the BIC to explore the number of genetic groups (K) that optimally describe the data.

Finally, we estimated genetic differentiation (F-statistics) and Analysis of Molecular Variance (AMOVA; Michalakis and Excoffier, 1996) assuming an infinite allele model where F-statistics correspond to those defined by Weir and Cockerham's θ (1984). F_{ST} measures the proportion of the total genetic variance contained within populations relative to the total genetic variance. This estimate is often used to assess gene flow between populations and can range between 0 and 1, where 0 indicates a completely panmictic/admixed population and 1 indicates complete isolation. To guide management considerations, we used thresholds proposed by Lowe and Allendorf (2010) as a method for gauging long-term genetic connectivity estimates among sites, where genetic connectivity is defined as "the degree to which gene flow affects evolutionary processes within subpopulations." They suggest that F_{ST} values > 0.2 represent too little genetic exchange between populations to reduce the harmful effects of genetic drift and inbreeding, which could lead to adaptive decline over time. The statistical significance of F_{ST} estimates was assessed with 10,000 permutations, and alpha significance ($\alpha = 0.05$) was adjusted for multiple tests using Bonferroni correction. We did not report alpha significance for cluster comparisons because the same data were used to infer the clusters (Meirmans, 2015).

Objective 3—Quantify Genetic Diversity

We calculated summary statistics in STACKS to compare genetic diversity among sites and genetic clusters across the lower Colorado River Basin. Summary statistics include the following: number of private alleles (P), mean nucleotide diversity (π), mean observed heterozygosity (H_{obs}), mean expected heterozygosity (H_{exp}, also known as gene diversity), and the inbreeding coefficient (F_{IS}). We also performed assessments of genetic diversity for United States populations (range-wide) relative to populations farther south along the Sierra Madre Occidental. We imported the variable call format (VCF) files, output from the PYRAD pipeline, into STACKS and computed the following summary statistics for each lineage identified from our phylogenetic analyses: number of variable single nucleotide polymorphism (SNP) loci (L_{Poly}), mean nucleotide diversity (π), and mean expected heterozygosity (H_{exp}).

We estimated the effective size (N_e) of each site as an additional measure of genetic diversity within the lower Colorado River Basin. Effective size (N_e) of a population is an important parameter for conservation management because it provides a way to quantify the amount of change in finite populations caused by genetic drift (chance loss of alleles through time) and inbreeding (Charlesworth, 2009; Hare and others, 2011). The value N_e also provides a measure of the ability of a population to maintain genetic diversity over generations to come (Wright, 1931). To this end, we relied on the N_e thresholds outlined in Frankham and others (2014) to help guide management considerations. They conclude that a minimum $N_e \geq 100$ as a short-term goal to avoid the risk of extinction owing to inbreeding depression, and $N_e \geq 1,000$ as a more long-term goal to maintain adaptive potential to environment change. We used the linkage disequilibrium method (LDNe; Waples and Do, 2008) within the program NEEESTIMATOR (Do and others, 2014) to obtain N_e values and parametric 95-percent confidence intervals (CI) for each site assuming random mating and a critical cut-off value for the frequency of rare alleles (*Pcrit*) of 0.02.

Finally, we investigated whether populations may have undergone population bottlenecks in the recent past (that is, 2Ne–4Ne generations; Luikart and Cornuet, 1998) using program bottleneck (Cornuet and Luikart, 1996). Populations that experience pronounced, recent reduction in the effective size (N_e) are predicted to show heterozygosity excess relative to the number of alleles, given that the number of alleles is expected to decline more quickly than heterozygosity in a declining population (Nei and others, 1975; Tajima, 1989). Because this method is sensitive to the mutational model under which the null range of alleles is simulated, we chose the infinite alleles model (IAM) for mutation (Haasl and Payseur, 2011) for SNPs and used 10,000 simulations and a Wilcoxon's signed rank test with alpha significance adjusted using Bonferroni correction.

Results

Thamnophis rufipunctatus

Summary of Bioinformatics

Using the PYRAD pipeline, our ddRAD sequencing effort yielded 503,673,323 total sequence fragments, with an average number of sequences per-individual at 2,632,299 (median: 2,228,887; min: 374,736; max: 15,831,632). The mean coverage depth per-individual was 65.3X (min: 25.4X; max: 144.8). After clustering and mindepth thresholds were applied, we had 88,816,567 total sites across 2,237,942 sequenced loci. We applied a filter that allowed 40 percent missing data per locus, which provided 5,606 homologous loci for phylogenetic analyses.

Using the STACKS pipeline, our ddRAD sequencing effort yielded an average of 3,000,204 sequences per-individual (median: 2,339,983; min: 390,948; max: 16,480,907). The mean coverage depth per-individual was 70.9X (min: 28.9X; max: 144.8X). After merging and calling final consensus sequences, we obtained 45,405 loci across 108 individuals sequenced. We obtained 2,505 loci after applying filters to retain all loci present across at least seven sampled sites and in 80 percent of the individuals sequenced with a minimum allele frequency cutoff set at 5 percent. We used STACKS to randomly select a single SNP from each of the 2,505 loci to generate a dataset consisting of 253 unlinked SNPs for population structure and genetic diversity analyses.

Objective 1—Phylogenetics Divergence of United States Populations

We used 79 individuals (plus two individuals of *T. eques* included as outgroups) across the Narrow-headed gartersnake species complex (*T. rufipunctatus*, *Thamnophis unilabialis*, and *Thamnophis nigronuchalis*) to test phylogenetic relationships inferred from Bayesian analysis. These analyses supported three major lineages with high posterior probability support ($Pp > 0.95$) and are consistent with the three geographically isolated species within the Narrow-headed gartersnake species complex (*T. rufipunctatus*, *T. unilabialis*, and *T. nigronuchalis*; fig. 4). Most importantly, we recovered a well-supported, reciprocally monophyletic lineage of *T. rufipunctatus* that was genetically divergent from the sister species *T. unilabialis*. In addition, the major clades within *T. rufipunctatus* were grouped according to their representative drainage basins (Verde, Salt, San Francisco, and Gila River Basins) across the lower Colorado River Basin. The only exception was the two samples from Tonto Creek (Salt River), which were grouped with samples from the Verde River instead of other Salt River sites. However, this inferred relationship was not strongly supported ($Pp < 0.95$; fig. 4).
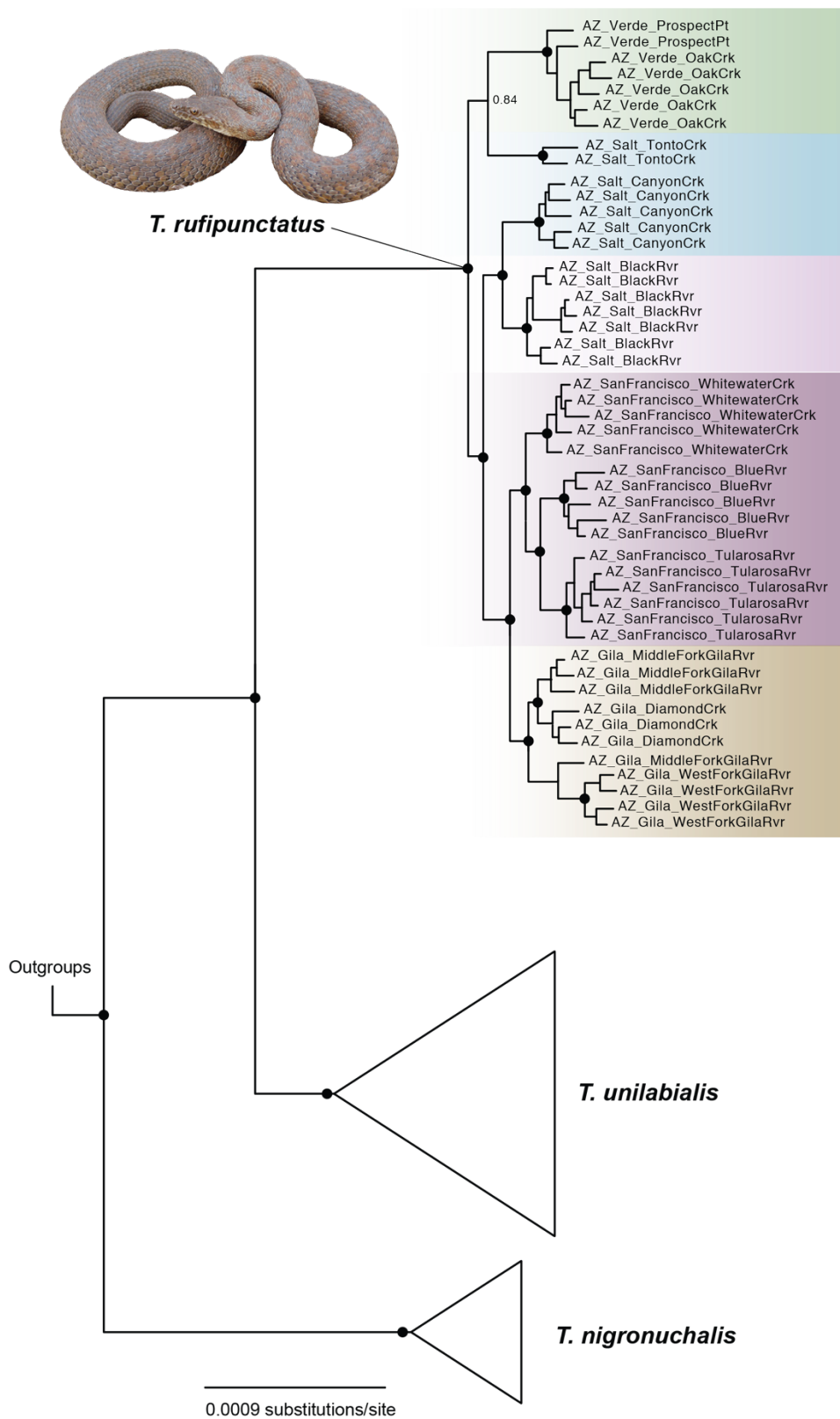
Figure 4. Phylogenetic tree of Narrow-headed gartersnake species complex based on 5,606 loci using MrBayes. The three species are delineated on each branch, and *Thamnophis unilabialis* and *Thamnophis nigronuchalis* clades were collapsed at the lineage node for clarity. Black dots indicate branch support of ≥ 0.95 posterior probability. The clades within *Thamnophis rufipunctatus* are colored according to population cluster analyses, and samples are labeled by "State_Basin_Site."

Objective 2—Spatial Patterns of Genetic Diversity

Thamnophis rufipunctatus exhibit significant population structure across their range, and the different methods of determining population structure produced similar results. First, Bayesian analysis in SNAPCLUST using the lowest BIC statistic partitioned *T. rufipunctatus* in to six major clusters (fig. 5a): a Verde River cluster (Oak Creek), two Salt River clusters (Canyon Creek and Black River), two San Francisco River clusters (Blue River and Tularosa River), and a Gila River cluster composed of Middle Fork Gi7la, Diamond Creek, and West Fork Gila sites. Snakes from the Saliz Creek site, which are present within the San Francisco River drainage, were also grouped within the Gila River cluster. However, these snakes were translocated from an admixed stock derived primarily from Middle Fork Gila and some individuals from Whitewater Creek (see appendix 1).

A
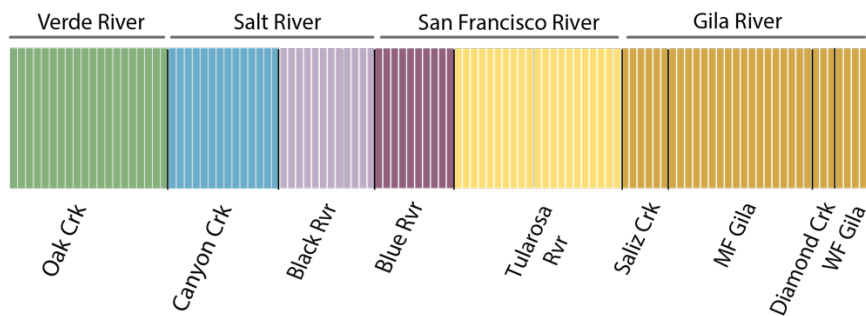
B
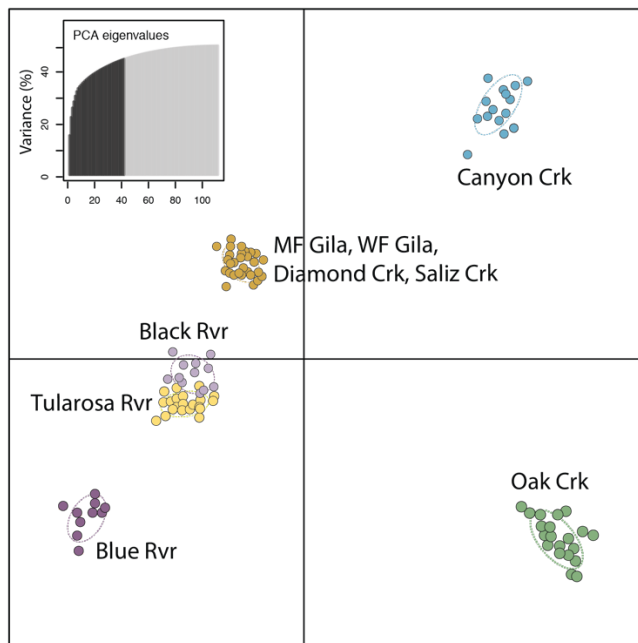
Figure 5. Cluster assignments of sampling sites for *Thamnophis rufipunctatus* using (a) SNAPCLUST, and (b) Discriminant Analysis Principal Components (DAPC). Shaded grey bars above (a) represent the major drainage basins within the lower Colorado River Basin.

12

In addition to SNAPCLUST, we used DAPC to explore genetic relationships among sites (fig. 5b). Although we obtained the lowest BIC score at $K = 8$, we explored DAPC among a range of K values (6–8) because the differences among the BIC scores were minimal. Once visualized using DAPC, it was evident that K values > 6 were no longer meaningful because no spatial differences were observed among additional clusters along the same axis. Similar results to the SNAPCLUST analysis were obtained using DAPC. Whereas most of the genetic clusters identified with DAPC occupy unique space along the axes, Black River and Tularosa River are grouped in similar space along one of the axes.

For most pairwise comparisons, F_{ST} estimates were high and exceeded limits for maintenance of genetic connectivity to avoid inbreeding or genetic drift among sites ($F_{ST} \geq 0.2$ in 26 of 36 comparisons; table 1). The global F_{ST} estimate was 0.391. For the seven sites with sample sizes greater than or equal to 6, all pairwise F_{ST} estimates were significant ($p < 0.001$ with Bonferroni correction; table 1). The lowest differentiation ($F_{ST} = 0.059$) was found between Middle Fork Gila and Diamond Creek, two adjacent localities within the Gila River drainage, whereas the highest level of differentiation ($F_{ST} = 0.431$) was found between Oak Creek and Tularosa River, located in the Verde and San Francisco River drainages, respectively.

Estimates of differentiation among the genetic clusters offered similar results (table 2). On the basis of cluster assignments, the lowest differentiation was found between the Tularosa and Gila River clusters ($F_{ST} = 0.147$) and highest differentiation between Verde (Oak Creek) and San Francisco (Tularosa River) River clusters ($F_{ST} = 0.429$). The results of the *a posteriori* AMOVA revealed that partitioning sites by clusters explained 37.4 percent of the molecular variance, whereas 34.0 percent of the molecular variance was accounted for when sites were partitioned by drainage basin (table 3).

Table 1. *Thamnophis rufipunctatus* pairwise genetic differentiation estimates (F_{ST}; Weir and Cockerham, 1984) for sampled sites.

[Statistical significance at $\alpha < 0.001$ after Bonferroni correction is indicated by bold face]

Sites	Oak Crk	Canyon Crk	Black Rvr	Blue Rvr	Tularosa Rvr	Saliz Crk	MF Gila	Diamond Crk	WF Gila
Oak Crk	—								
Canyon Crk	**0.364**	—							
Black Rvr	**0.395**	**0.198**	—						
Blue Rvr	**0.411**	**0.321**	**0.239**	—					
Tularosa Rvr	**0.431**	**0.334**	**0.253**	**0.158**	—				
Saliz Crk	**0.384**	**0.281**	**0.232**	**0.175**	**0.162**	—			
MF Gila	**0.336**	**0.301**	**0.276**	**0.266**	**0.226**	**0.111**	—		
Diamond Crk	0.427	0.343	0.289	0.231	0.202	0.136	0.059	—	
WF Gila	0.422	0.338	0.298	0.215	0.222	0.128	0.093	0.162	—

Table 2. *Thamnophis rufipunctatus* pairwise genetic differentiation estimates (F_{ST}; Weir and Cockerham, 1984) for clusters assignments.

[Statistical significance was not assessed for pairwise comparisons because the same data were used to infer the cluster assignments (Meirmans, 2015)]

K	1	2	3	4	5	6	Clusters
1	—	—	—	—	—	—	Oak Creek
2	0.358	—	—	—	—	—	Canyon Creek
3	0.399	0.211	—	—	—	—	Black River
4	0.397	0.297	0.231	—	—	—	Blue River
5	0.429	0.317	0.257	0.169	—	—	Tularosa River
6	0.338	0.257	0.183	0.178	0.147	—	Gila River (MF & WF Gila, Diamond Crk)

Table 3. Analysis of molecular variance (AMOVA) comparisons among different partitioning strategies.

Source of genetic variance	Percent variation	
	Drainage basin *K*=4	Clusters *K*=6
Within sites	53.4%	55.6%
Among sites within clusters	12.7%	0.07%
Among clusters	34.0%	37.4%

Objective 3—Genetic Diversity

Comparisons of genetic diversity revealed lower diversity in lower Colorado River Basin populations relative to sister species populations along the Sierra Madre Occidental (table 4), with Sierra Madre Occidental sister species (*T. unilabialis*) harboring between 1.7 to 2.9 times as much diversity as *T. rufipunctatus* depending on the diversity statistic used.

Table 4. Comparisons of genetic diversity of *Thamnophis rufipunctatus* and *Thamnophis eques megalops* relative to their sister species/sister lineages found along the Mexican highland.

[Abbreviations and indices are as follows: LCRB, lower Colorado River Basin; nSMO, northern Sierra Madre Occidental; sSMO, southern Sierra Madre Occidental; TVB, Trans-Volcanic Belt; N, sample size; L_{Poly}, number of single nucleotide polymorphisms (SNPs); π, mean nucleotide diversity; H_{exp}, mean expected heterozygosity (gene diversity)]

Sister Species / Populations	Range	N	L_{Poly}	π	H_{exp}
T. rufipunctatus	LCRB	32	330	0.053	0.055
T. unilabialis	nSMO	22	940	0.152	0.148
T. nigronuchalis	sSMO	4	157	0.109	0.096
T. e. megalops	LCRB	49	250	0.038	0.038
T. e. megalops	nSMO	19	236	0.054	0.052
T. e. megalops	sSMO	18	280	0.078	0.076
T. eques	TVB	16	411	0.090	0.087

Genetic diversity for *T. rufipunctatus* sites and clusters are given in table 5 and visualized for sampled sites in figure 6. Both Oak and Canyon Creeks had much greater numbers of private alleles than all other sites sampled. Mean nucleotide diversity and heterozygosity were generally similar across sites. Inbreeding coefficients (F_{IS}) ranged from −0.093 to 0.047, with the highest values observed at Oak Creek, Blue River, Tularosa River, and Middle Fork Gila (see appendix 2). The highest diversity and lowest F_{IS} estimates were both observed at Saliz Creek, consistent with the admixed nature of the site (see appendix 1). Genetic diversity within the clusters followed similar patterns as sampled sites.

Table 5. Genetic diversity of *Thamnophis rufipunctatus* for sampling locations (based on 253 single nucleotide polymorphisms, SNPs) and clusters (based on 265 SNPs) within the lower Colorado River Basin.

[Indices are as follows: N, sample size; P, number of private alleles; π, mean nucleotide diversity; H_{obs}, mean observed heterozygosity; H_{exp}, mean expected heterozygosity (gene diversity); F_{IS}, inbreeding coefficient; Bottleneck, P-values for Wilcoxon tests for heterozygosity excess; —, indicates indices were not computed. Using Bonferroni correction, alpha significance for Bottleneck was assessed at $P < 0.0083$ and significant values are in bold face]

Groupings	N	P	π	H_{obs}	H_{exp}	F_{IS}	Bottleneck
Oak Creek	20	43	0.148	0.118	0.127	0.034	**0.0000**
Canyon Creek	14	21	0.167	0.145	0.140	0.001	**0.0000**
Black River	12	1	0.149	0.134	0.134	0.019	**0.0001**
Blue River	10	3	0.190	0.165	0.171	0.038	**0.0000**
Tularosa River	21	3	0.145	0.114	0.124	0.033	**0.0000**
Saliz Creek	6	1	0.224	0.249	0.182	−0.093	—
Middle Fork Gila	18	0	0.152	0.118	0.132	0.047	**0.0030**
Diamond Creek	3	0	0.124	0.114	0.099	0.011	—
West Fork Gila	4	0	0.181	0.153	0.140	0.016	—
Clusters (*K*=6)							
Oak Creek	20	40	0.156	0.121	0.135	—	—
Canyon Creek	14	19	0.182	0.169	0.162	—	—
Black River	12	1	0.143	0.128	0.130	—	—
Blue River	10	4	0.192	0.168	0.172	—	—
Tularosa River	21	2	0.137	0.111	0.116	—	—
Gila River	25	6	0.176	0.137	0.153	—	—
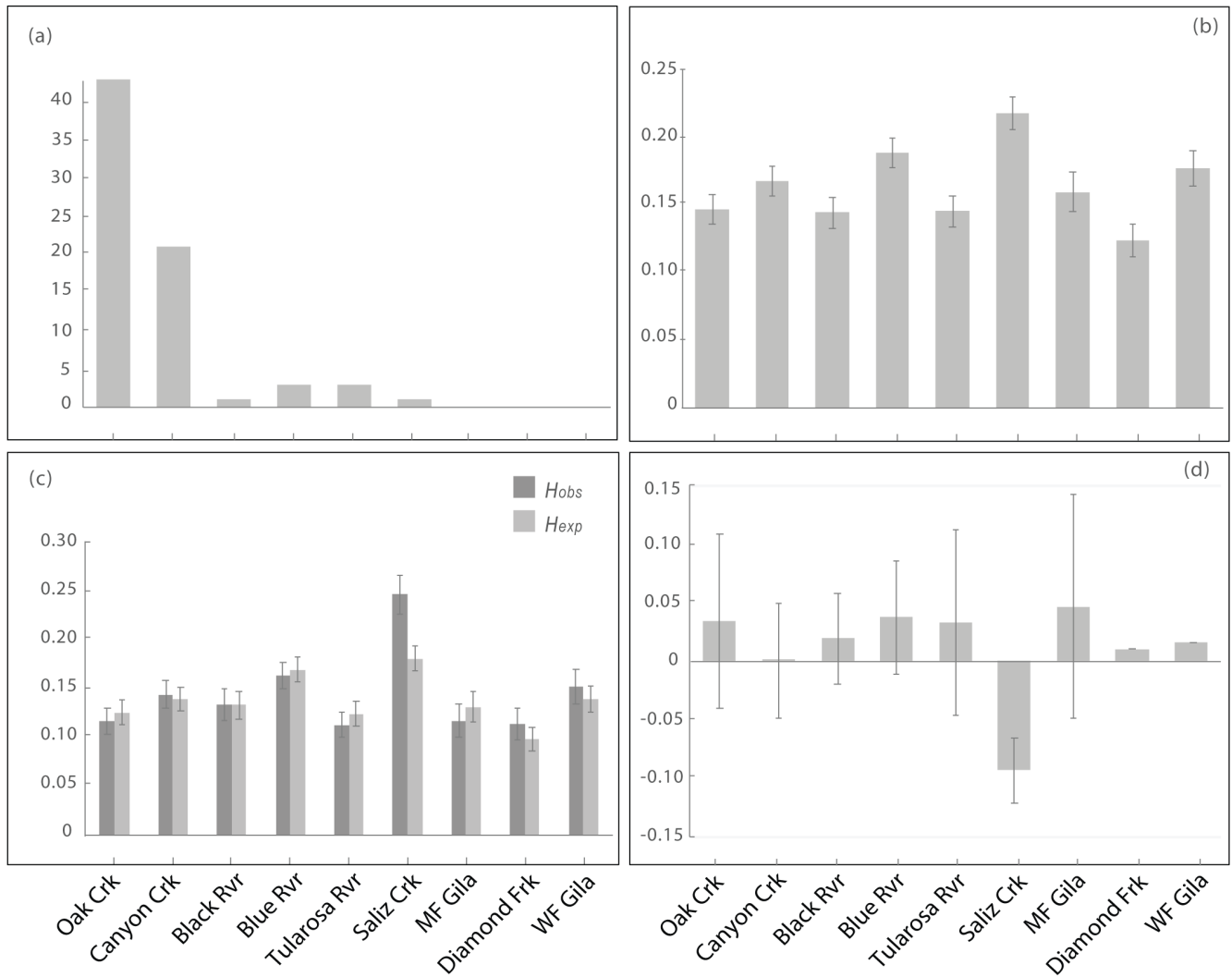
Figure 6. Genetic diversity summary statistics at sampling sites for *Thamnophis rufipunctatus* with standard errors shown around each mean: (a) number of private alleles, *P*; (b) nucleotide diversity, π; (c) observed, H_{obs}, and expected heterozygosity, H_{exp}; and (d) inbreeding coefficients, F_{IS}.

Estimates of effective size (N_e) across sites resulted in values well below the short-term threshold recommendation to limit inbreeding depression (≥ 100), with point estimates ranging from 13 to 42 (table 6). Effective sizes were particularly low at Canyon Creek, Blue River, and Middle Fork Gila. Consistent with these low effective population size estimates, we detected significant bottlenecks at all sites tested (table 5), suggesting a loss of genetic diversity within the last 2–4 generations has occurred across the range of this species.

16

Table 6. Estimates of effective size and parametric 95-percent confidence intervals across sites of *Thamnophis rufipunctatus* for sampling locations with greater than 10 samples using the linkage disequilibrium method.

[Abbreviations are as follows: N_e, effective size of the population; CI, confidence interval; Inf, infinite and represents an undefined upper limit; —, indicates sites that we were not able to obtain an estimate]

Groupings	N	Effective size (N_e)	Lower 95% CI	Upper 95% CI
Oak Creek	20	24	17	37
Canyon Creek	14	19	14	26
Black River	12	42	23	160
Blue River	10	16	11	23
Tularosa River	21	30	22	46
Middle Fork Gila	18	13	8	21

Thamnophis eques megalops

Summary of Bioinformatics

Using the PYRAD pipeline, our ddRAD sequencing effort yielded 323,215,640 total sequence fragments, with an average number of sequences per-individual at 2,032,803 (median: 1,633,989; min: 192,368; max: 7,165,848). The mean coverage depth per-individual was 40X (median: 33.6X, minimum: 12.3X, maximum: 111.1X). After clustering and mindepth filtering thresholds were applied, we had 57,729,972 total sites across 1,456,144 sequenced loci. We applied a filter that allowed 30 percent missing data per locus, which provided 2,278 homologous loci for phylogenetic analyses.

Using the STACKS pipeline, our ddRAD sequencing effort yielded an average of 2,256,163 sequences per-individual (median: 1,883,068; min: 277,240; max: 7,368,397). The mean coverage depth per-individual was 50.2X (min: 25.9X; max: 93.4X). After merging stacks and calling final consensus sequences, we obtained 202,971 loci across 132 individuals sequenced. We obtained 2,948 loci after applying filters to retain all loci present across at least seven sampled sites and in 80 percent of the individuals sequenced with a minimum allele frequency cut off set at 5 percent. We used STACKS to randomly select a single SNP from each of the 2,948 loci to generate a dataset consisting of 537 SNPs for population structure and genetic diversity analyses.

Objective 1—Phylogenetics Divergence of United States Populations

The phylogenetic relationships inferred from Bayesian analysis of 2,278 concatenated RADseq loci sequenced across 106 individuals supported four major lineages with high posterior probabilities (PP > 0.95) that subdivide the range of *T. eques* in to a lower Colorado River Basin lineage, a northern Sierra Madre Occidental lineage, a southern Sierra Madre Occidental lineage, and a Transvolcanic Belt lineage (fig. 7). Notably, this analysis indicates that United States populations of *T. e. megalops* are genetically divergent peripheral isolates that represent a distinct lineage from populations inhabiting higher elevations along the Sierra Madre Occidental. Each of the major clades within the lower Colorado River Basin was grouped according to their representative drainage basins (Bill Williams, Verde, Salt, Santa Cruz, and Gila River Basins).
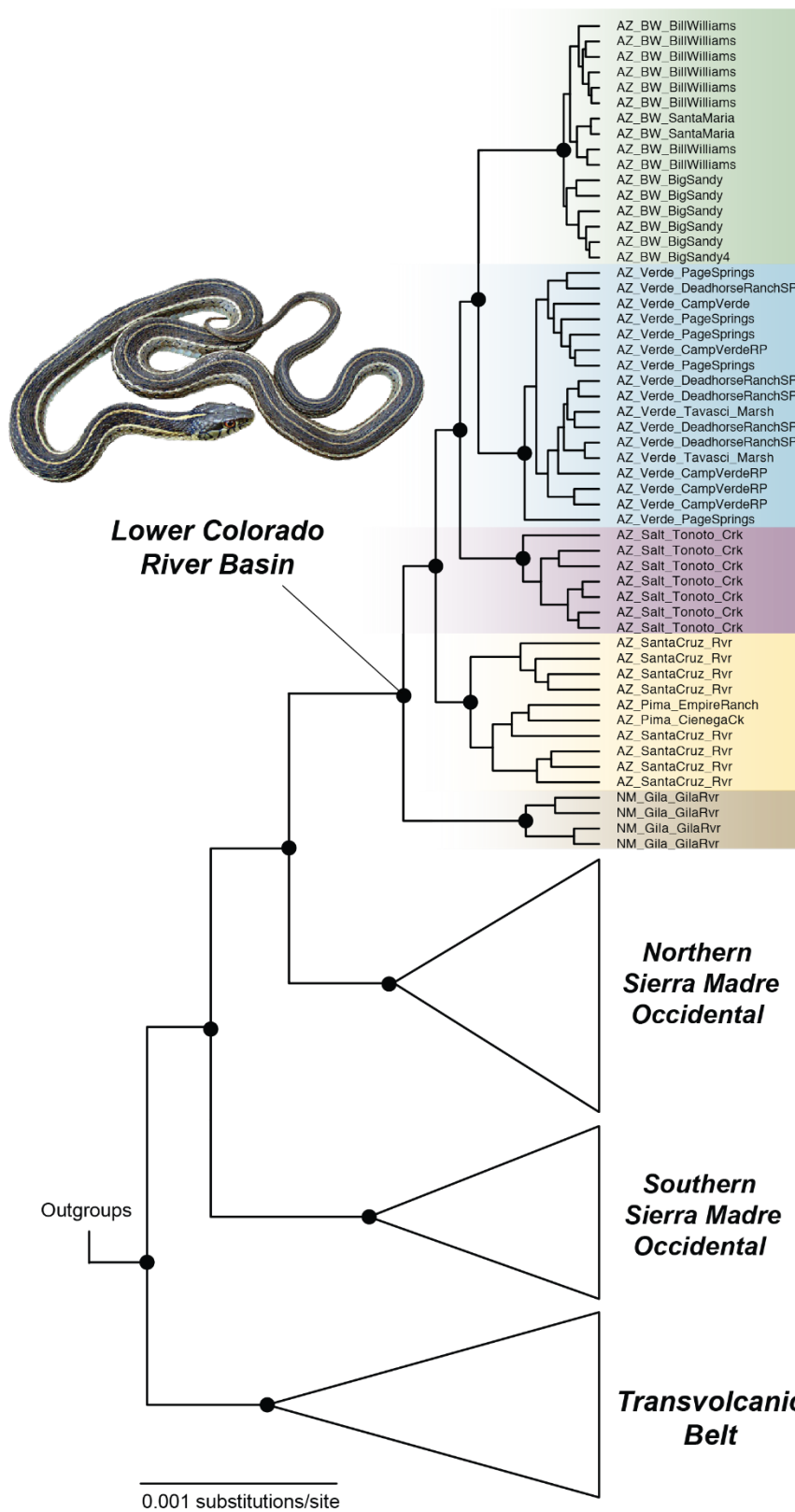
Figure 7. Phylogenetic tree of Northern Mexican gartersnake (*Thamnophis eques megalops*) based on 2,278 loci using BEAST. The clades for each lineage along the Mexican highlands are collapsed at the node. Black dots indicate branch support of ≥0.95 posterior probability. The clades within lower Colorado River Basin are colored according to Bayesian population cluster analyses, and samples are labeled by "State_Basin_Site."

Objective 2—Spatial Patterns of Genetic Structure

Thamnophis eques megalops exhibited significant population structure across the 10 United States sites evaluated, and the different methods of determining population structure produced the same results. First, maximum likelihood cluster analysis in SNAPCLUST using the BIC statistic partitioned *T. e. megalops* in to five major clusters that group genetic variation according to major drainage basins (fig. 8a). The five clusters include the Bill Williams River (Centennial Wash, Santa Maria, and Big Sandy), Verde River (Camp Verde, Dead Horse Ranch, Page Springs), Salt River (Tonto Creek), Santa Cruz River (San Rafael Valley), and Gila River. We also used DAPC to explore cluster patterns among sites. We obtained the lowest BIC score at *K*=5, and DAPC offers similar results to the SNAPCLUST analysis (fig. 8b).
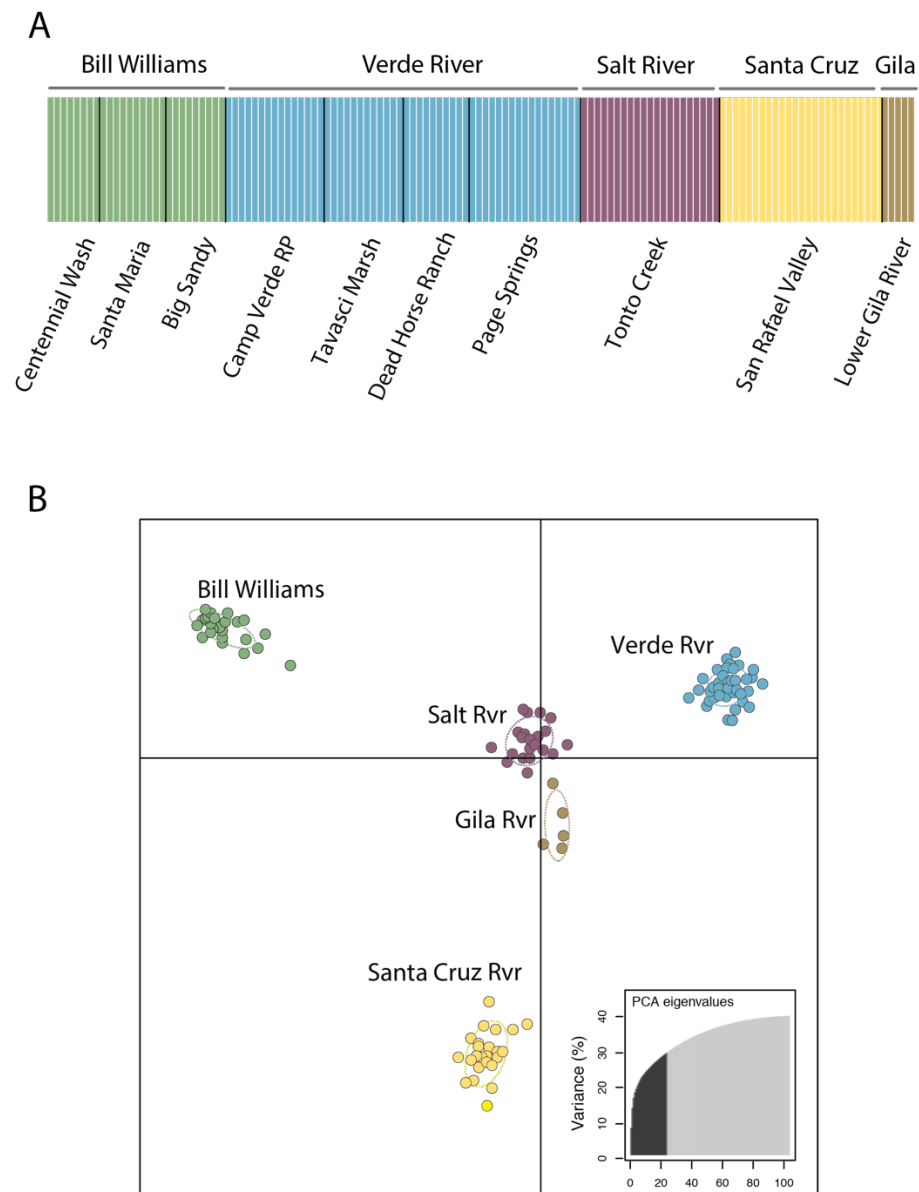
Figure 8. Cluster assignments of sampling sites for *Thamnophis eques megalops* using (a) SNAPCLUST and (b) Discriminant Analysis Principal Components (DAPC). Shaded grey bars above (a) represent the major drainage basins.

The global F_{ST} estimate was 0.286. Pairwise F_{ST} comparisons of sites within the same drainage basin (for example, Bill Williams and Verde River Basin sites) fell within the threshold limits for maintenance of genetic connectivity to avoid inbreeding or drift among sites ($F_{ST} \geq 0.2$ in 0 of 9 pairwise comparisons; table 7). However, F_{ST} for most pairwise comparisons across drainage basins exceeded these limits ($F_{ST} \geq 0.2$ in 25 of 45 pairwise comparisons; table 7). Across the 10 sites, pairwise F_{ST} estimates were significant (p < 0.001 with Bonferroni correction); the only exceptions were between sampled sites within the Verde River (Tavasci Marsh and Dead Horse Ranch). Estimates of differentiation among the genetic clusters were high, ranging from 0.121 (lowest) between the Santa Cruz and Gila Rivers to 0.361 (highest) between Bill Williams and Gila Rivers, the farthest separated drainage basins (table 8). For *T. e. megalops*, partitioning sites by clusters was the same as separating sites by major drainage basins. The results of the *a posteriori* AMOVA revealed that the five clusters explained 36.1 percent of the molecular variance, with 8.8 percent explained by among populations within clusters, and 55.1 percent explained by within-population variation.

Table 7. *Thamnophis eques megalops* pairwise genetic differentiation estimates (F_{ST}; Weir and Cockerham 1984) for sampled sites.

[Statistical significance at α < 0.001 after Bonferroni correction is indicated by bold face]

Sites	Centennial Wash	Santa Maria	Big Sandy	Camp Verde	Tavasci Marsh	Dead Horse Ranch	Page Springs	Tonto Creek	San Rafael	Gila River
Centennial Wash	—									
Santa Maria	**0.099**	—								
Big Sandy	**0.100**	**0.099**	—							
Camp Verde	**0.283**	**0.348**	**0.320**	—						
Tavasci Marsh	**0.289**	**0.339**	**0.339**	**0.059**	—					
Dead Horse Ranch	**0.296**	**0.347**	**0.337**	**0.053**	0.036	—				
Page Springs	**0.305**	**0.346**	**0.352**	**0.057**	**0.067**	**0.066**	—			
Tonto Creek	**0.209**	**0.242**	**0.234**	**0.123**	**0.120**	**0.117**	**0.148**	—		
San Rafael	**0.209**	**0.234**	**0.236**	**0.179**	**0.168**	**0.159**	**0.191**	**0.162**	—	
Gila River	**0.209**	**0.453**	**0.448**	**0.263**	0.288	**0.290**	**0.306**	**0.199**	**0.122**	—

Table 8. *Thamnophis eques megalops* pairwise genetic differentiation estimates (F_{ST}; Weir and Cockerham 1984) for clusters assignments.

K	1	2	3	4	5	Clusters
1	—	—	—	—	—	Bill Williams (Centennial, Santa Maria, Big Sandy)
2	0.231	—	—	—	—	Verde River (Camp Verde, Tavasci, Marsh, Deadhorse Ranch, Page Springs)
3	0.256	0.134	—	—	—	Salt River (Tonto Creek)
4	0.246	0.182	0.164	—	—	Santa Cruz River (San Rafael Valley)
5	0.361	0.202	0.191	0.121	—	Gila River (New Mexico)

Objective 3—Genetic Diversity

Comparisons of *T. e. megalops* genetic diversity (H_S) revealed lower diversity in United States populations relative to sister populations along the Sierra Madre Occidental (table 4), with Sierra Madre Occidental sister populations harboring between 1.4 and 2.4 times as much diversity as United States populations. Estimates of inbreeding coefficients (F_{IS}) were not significantly different among United States and Sierra Madre Occidental populations.

Genetic diversity for sites and clusters is given in table 9 and visualized for sampled sites in figure 9. The number of private alleles in San Rafael Valley was much greater (60) than all other sites ($P \leq 13$). Overall, genetic diversity indices varied longitudinally and were highest within the eastern sites and lowest in western sites (fig. 9). Inbreeding coefficients (F_{IS}) ranged from −0.042 to 0.099, with highest coefficients estimated at Camp Verde Regional Park, Tonto Creek, and San Rafael Valley.

Table 9. Genetic diversity of *Thamnophis eques megalops* for sampling locations (based on 537 single nucleotide polymorphisms, SNPs) and clusters (based on 252 SNPs) within the lower Colorado River Basin.

[Indices are as follows: N, sample size; P, number of private alleles; π, mean nucleotide diversity; H_{obs}, mean observed heterozygosity; H_{exp}, mean expected heterozygosity (gene diversity); F_{IS}, inbreeding coefficient; Bottleneck, P-values for Wilcoxon tests for heterozygosity excess; —, indicates that indices were not computed. Using Bonferroni correction, alpha significance for Bottleneck was assessed at $P < 0.0056$ and significant values are in bold face]

Groupings	N	P	π	H_{obs}	H_{exp}	F_{IS}	Bottleneck
Centennial Wash	8	0	0.080	0.093	0.067	–0.042	0.2918
Santa Maria	10	3	0.079	0.075	0.064	–0.011	**0.0000**
Big Sandy	9	0	0.074	0.061	0.057	0.001	0.0092
Camp Verde	15	1	0.178	0.153	0.162	0.050	**0.0001**
Tavasci Marsh	12	0	0.168	0.155	0.155	0.024	**0.0006**
Dead Horse Ranch	10	1	0.173	0.160	0.158	0.022	**0.0006**
Page Springs	17	0	0.156	0.147	0.142	0.008	**0.0000**
Tonto Creek	21	13	0.223	0.198	0.217	0.074	**0.0000**
San Rafael Valley	25	58	0.287	0.252	0.282	0.099	**0.0000**
Gila River	5	5	0.187	0.167	0.159	0.033	—
Clusters (*K*=5)							
Bill Williams	27	6	0.085	0.071	0.084	—	—
Verde River	54	7	0.166	0.141	0.160	—	—
Salt River	21	5	0.222	0.196	0.217	—	—
Santa Cruz River	25	20	0.273	0.241	0.267	—	—
Gila River	5	0	0.180	0.179	0.163	—	—
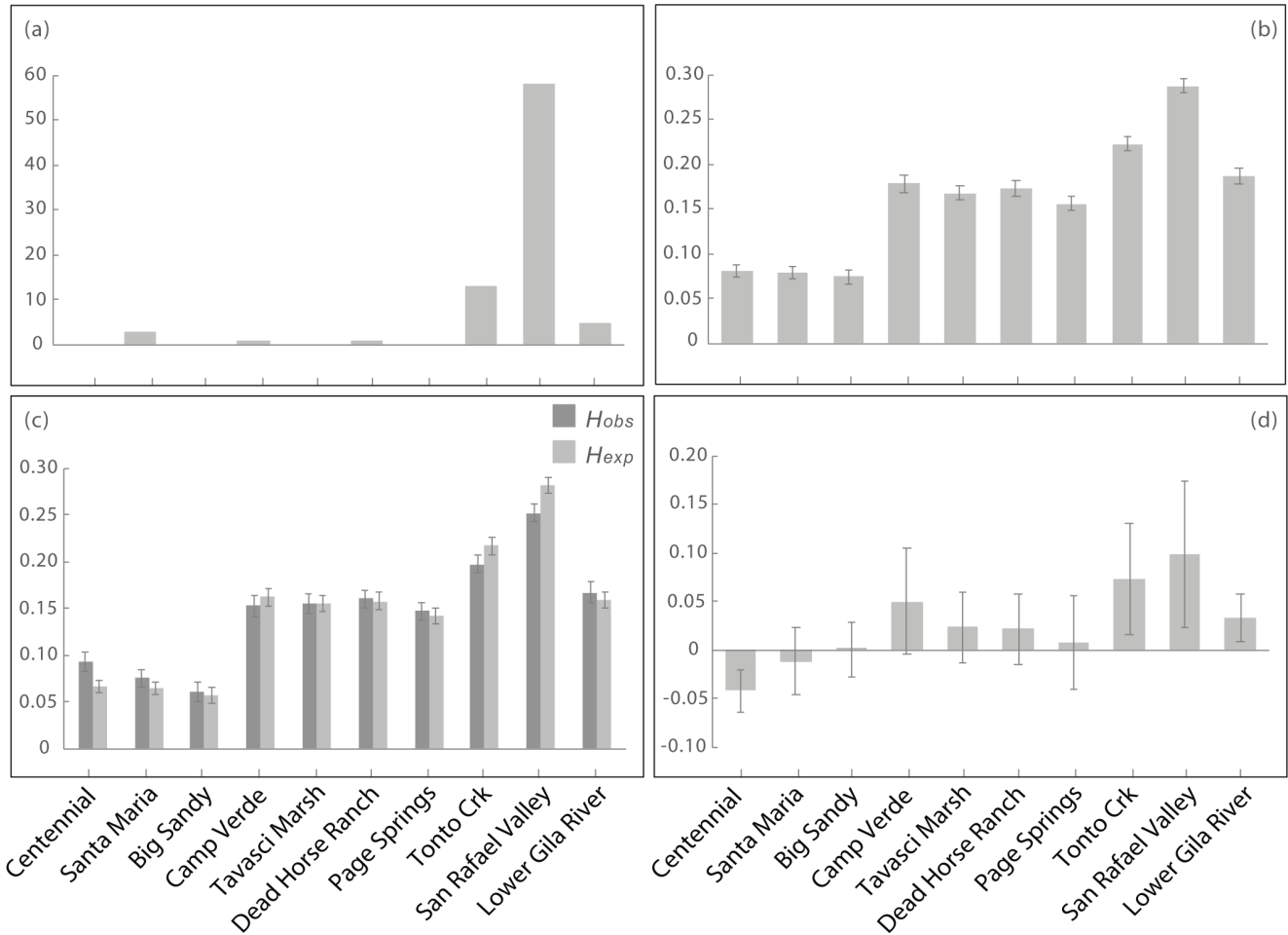
Figure 9. Genetic diversity summary statistics at sampling sites for *Thamnophis eques megalops* with standard errors shown around each mean: (a) number of private alleles, *P*; (b) nucleotide diversity, π; (c) observed, H_{obs}, and expected heterozygosity, H_{exp}; and (d) inbreeding coefficients, F_{IS}.

Estimates of effective population size (N_e) across sites resulted in a broad range of values (from 15 to 204, table 10), with six of the seven estimates falling below the recommended N_e threshold to preclude inbreeding depression (≥ 100); only San Rafael Valley exceeded this threshold. Effective population sizes were especially low at Big Sandy and Tonto Creek ($N_e \sim 15$). Despite overall higher effective sizes (compared to *T. rufipunctatus*), we detected evidence for significant bottlenecks at 7 of 9 sites tested (table 9), indicating a loss of genetic diversity has occurred within the last 2–4 generations across these sites.

Table 10. Estimates of effective size (N_e) and parametric 95-percent confidence intervals across sites of *Thamnophis eques megalops* for sampling locations.

[Abbreviations are as follows: N_e, effective size of the population; CI, confidence interval; Inf, infinite and represents an undefined upper limit; —, indicates sites that we were not able to obtain an estimate]

Groupings	N	Effective size (N_e)	Lower 95% CI	Upper 95% CI
Centennial Wash	8	—	—	—
Santa Maria	10	63.4	28.8	Inf
Big Sandy	9	15.8	7.8	58
Camp Verde	15	73.6	49.8	135.2
Tavasci Marsh	12	—	—	—
Dead Horse Ranch SP	10	81.5	42.2	Inf
Page Springs	17	63.5	47.2	94.8
Tonto Creek	21	15.1	14.2	16.1
San Rafael Valley	25	204	143.6	349.9

Discussion

Our work represents the most extensive genetic study of these threatened gartersnake species undertaken to date. Using genome-wide markers, we evaluated the population structure and quantified genetic variation of *Thamnophis rufipunctatus* and *T. eques megalops*. Our results show that both species have highly structured populations across the lower Colorado River Basin in the United States and that these populations are differentiated from sister species/populations located farther south in Mexico. In general, both species exhibit comparably reduced genetic diversity across the lower Colorado River Basin relative to their counterparts along the Sierra Madre Occidental in Mexico and have small effective population sizes. We discuss these patterns in more detail for each species and suggest how these results might be used to inform conservation management.

Phylogenetic Divergence and Diversity of Lower Colorado Basin and Sierra Madre Occidental Populations

On the basis of over a thousand RADseq markers, our phylogenetic analyses corroborate past findings for *T. rufipunctatus* and represent novel results for *T. e. megalops*. Specifically, populations along the lower Colorado River Basin represent genetically divergent lineages separate from their more southern sister species (*T. unilabialis*) or sister populations (other *T. e. megalops* populations) that inhabit the Mexican highlands of the Sierra Madre Occidental. Wood and others (2011) suggested a paleo-hydrological connection had formerly linked river drainages of the lower Colorado River Basin with more southern drainage basins in Mexico and estimated the timing of lineage divergence between *T. rufipunctatus* (in the United States) and *T. unilabialis* (in northern Mexico) within the late Pleistocene. During this time frame, the region experienced repeated climatic shifts causing regular patterns of connectedness and isolation of aquatic systems that eventually severed the hydrological connection and led to the large distributional gap between *T. rufipunctatus* and *T. unilabialis* (fig. 3a). Given the similar ecological affinities and geographic distributions of *T. e. megalops*, our results indicate that related historical climatic events may have given rise to the lineage divergence observed in *T. e. megalops* between the lower Colorado River Basin and Sierra Madre Occidental populations. Although we were not able to secure samples of *T. e. megalops* from intermediary populations scattered along the lower elevation foothills of eastern Sonora, Mexico, these populations are geographically isolated from higher elevation populations in the Sierra Madre Occidental, and studies of fishes occupying the eastern Sonora river systems have generally shown closer genetic relationships to Gila River drainages within the lower Colorado River Basin (Smith and Miller, 1986; Meffe and Vrijenhoek, 1988; Schönhuth and others, 2011). Efforts to acquire and sequence *T. e. megalops* samples from this region would help to identify their genetic affinity and may be useful information for conservation of this species, if management strategies become more reliant on conservation translocation and genetic rescue measures (see below).

Population Structure and Diversity Within the Lower Colorado River Basin

For both species, comparisons of genetic variation indicated that lower Colorado River Basin populations in the United States harbor lower levels of genetic diversity than each of their closely related sister species/populations in northern Sierra Madre Occidental. This may be due, in part, to the smaller distribution ranges for both species of gartersnakes within the lower Colorado River Basin compared to the northern Sierra Madre Occidental ranges of their counterparts, but the accelerated anthropogenic disturbances throughout the lower Colorado River Basin (habitat loss and degradation, competition and predation by non-native species) relative to those along Sierra Madre Occidental may also be an important contributing factor. The lower diversity estimates recovered in *T. rufipunctatus* are especially noteworthy, as its entire range is in the lower Colorado River Basin and remaining populations are small and isolated.

Throughout the lower Colorado River Basin, our results reveal highly structured populations for both species of gartersnakes that point to shared contemporary and historical drivers of differentiation. This high differentiation may reflect increased insularization of formerly panmictic distributions owing to the severe abiotic conditions and natural habitat fragmentation that has persisted in this region over the last 10,000- to 12,000-year post-pluvial period (Axelrod, 1979; Morrison, 1991; Metcalfe, 2006). Alternatively, populations throughout the lower Colorado River Basin may have experienced higher connectivity in the past, but more recent drift-mediated processes may have driven the high differentiation owing to the loss of physical and genetic connectivity within drainage basins as a result of recent anthropogenic disturbances. We note that high differentiation (F_{ST}) within and among drainage basins was also coupled with population bottlenecks and small effective sizes at most sites. Population surveys over the past several decades have provided strong evidence of population retractions and declines in abundance at occupied sites (Rosen and Schwalbe, 1988; Rosen, 1993; Nowak and Santana-Bendix, 2002; Holycross and others, 2006; Nowak, 2006; Hibbitts and others, 2009; Jennings and Christman, 2015; Nowak and others, 2017), which indicates that anthropogenic disturbances also drive loss of connectivity and may have purged populations of alleles that were once shared among drainages. However, disentangling these different temporal influences on genetic differentiation and establishing a direct link to more recent anthropogenic disturbance is challenging, requiring evaluation of historical genetic samples taken from time intervals closer to pre-human interferences.

The strongest patterns of genetic variation partitioning that was supported by phylogenetic, genetic clustering, and AMOVA analyses were largely within hydrologic drainage basins for both species, supporting a hierarchical basin structure with varying degrees of connection and isolation operating largely as a function of geographic distance. For *T. rufipunctatus,* lingering signals of past connectivity between the Verde and Salt River Basins may be evidenced by the sister clade relationship observed between the Oak Creek and Tonto Creek samples in the phylogenetic analysis (fig. 4). However, this interpretation of genetic connectivity should be held as tentative, given the low posterior support and small sample size at Tonto Creek (n = 2). Patterns of former connectivity among the five major drainages occupied by *T. e. megalops* was less clear; we found no evidence of admixed coancestry across drainage basins. To our knowledge, the sites we sampled represent the major extant populations within the lower Colorado River Basin distribution, and the large geographic distance that now separates these populations may be why we recovered such strong evidence of genetic isolation. Considering the current spatial and genetic structure of both species, along with the host of ecological and landscape disturbances, migrant exchange within and among drainages is probably unlikely.

Applying Results to Population-Level Conservation Management

On the basis of population genetic theory, an ideal, genetically healthy population is one that has a sufficiently large effective population size (N_e) with moderate connectivity with sister populations to maintain evolvability and counter the effects of genetic drift and inbreeding depression. When connectivity across the landscape is disrupted, the levels and distribution of genetic diversity can increasingly erode, leading to increased genetic differentiation between populations and small effective population sizes (Lowe and Allendorf, 2010). As N_e decreases, a synergistic feedback between genetic drift and inbreeding can lead to decreased fitness of a population (inbreeding depression) and increased susceptibility to environmental and demographic changes (Miller and Hobbs, 2002; Frankham, 2005; Fischer and Lindenmayer, 2007). These changes can drive further demographic declines and push already threatened populations closer to extirpation (Gilpin and Soulé, 1986). Our work suggests that many sites, across both species, have low genetic diversity as indicated by the low effective population sizes, signatures of genetic bottlenecks, and in some cases, high inbreeding coefficients (see appendix 2). These signals are more pronounced in *T. rufipunctatus*. For example, estimates of effective size (N_e) in *T. rufipunctatus* were well below the short-term threshold recommendation that would limit inbreeding depression ($N_e \geq 100$), and these estimates were coupled with bottlenecks. Effective size estimates for *T. e. megalops* were generally higher across sites but were similarly below 100 except at San Rafael Valley. Sites with low effective size that lack genetic diversity may have limited capacity to adapt to rapid environmental changes. Furthermore, we found high F_{ST} estimates that approached or exceeded limits for maintenance of genetic connectivity to reduce the negative effects of drift and inbreeding among sites in both species (but more so for *T. rufipunctatus*). This highlights the potential role of life history (for example, local abundance, reproductive variance, dispersal abilities) and natural history traits (for example, restricted diet and habitat use of *T. rufipunctatus* compared to *T. e. megalops*) in determining how genetic erosion progresses. Research aimed at unravelling these differences may help explain why *T. rufipunctatus* appears more vulnerable to local extinctions. Taken together, the effects of genetic drift and continued loss of genetic variation may be a serious threat to population persistence in these small, isolated gartersnake populations, especially in the face of future climate change (Giermakowski and others, 2015).

When more contemporary disturbances are the likely driver of non-natural genetic differentiation, conservation strategies may focus on reinstating connectivity among populations. For threatened and endangered species, recovery goals generally focus on ameliorating ecological threats and restoring habitat to ensure self-sustaining populations (Doak and others, 2015). However, genetic management can also be incorporated into conservation strategies, given that population persistence is generally positively linked to genetic diversity (Frankham and Ralls, 1998; Pierson and others, 2016). Assisted migration has been shown to be useful for management of small, at-risk populations (Weeks and others, 2011; Frankham, 2015; Whiteley and others, 2015; Ralls and others, 2017). Many recent studies have shown that augmentation (introducing a limited number of known unrelated and disease-free immigrants to threatened populations) can have favorable population-level fitness responses (increased genetic variation, fecundity, and population growth) that far outweigh concerns of outbreeding depression (Madsen and others, 2004; Johnson and others, 2010; McEachern and others, 2011; Miller and others, 2012; Smyser and others, 2013; Fitzpatrick and others, 2016; Weeks and others, 2017). The alternative approach, no assisted migration, may place populations at greater risk to many interacting stochastic processes that will continue to erode genetic diversity and compromise their capacity to adapt to changing environmental conditions.

Should a genetic management and recovery strategy be adopted, the population genetic results provided here help highlight sites of concern and locate the most genetically similar sites for translocation and repatriation efforts. This information can be used in concert with established captive husbandry research programs at the Phoenix Zoo and Northern Arizona University. Such measures could help curb further population genetic change, alleviate problems associated with low genetic diversity, and strengthen fitness and (or) adaptive potential across the range of these two gartersnake species. Although relationships between fitness and genetic diversity have not been directly measured in these species, data from *T. rufipunctatus* at Saliz Creek, a translocation site inadvertently comprised of snakes from different sources (appendix 1), showed increased genetic diversity and lower inbreeding coefficients relative to the source populations. These preliminary results indicate that genetic management could be a useful option, and continued monitoring at Saliz Creek may provide valuable information to better assess the possible benefits of genetic management and success of translocations. We emphasize that a genetic management strategy may work best when coupled with other species-specific management actions designed to reduce threats and improve demographic viability and increase carrying capacities at individual sites (Frankham and others, 2017; Love Stowell and others, 2017; Ralls and others, 2017). Prior to implementing assisted gene flow measures, it is important to assess the risks of inbreeding versus outbreeding and potential loss of local adaptation for admixed populations (Hedrick and Fredrickson, 2010). However, the risk of outbreeding depression or loss of adaptation is very low between recently diverged populations, especially when that divergence is primarily caused by anthropogenic barriers (Frankham and others, 2011). For both gartersnake species, we suggest separate management of different basins (with some exceptions, see below) as a conservative option until an informed risk-benefit analysis for specific populations suggests otherwise (example work flows can be found in Frankham and others, 2011; Ralls and others, 2017). Risks could include introducing disease or pathogen agents or reduction in local fitness if adaptive differences exist. Managing each drainage basin as separate genetic units is consistent with population structure detected from phylogenetic and cluster analyses and could be beneficial for long-term population viability. Specifically, we suggest reconnecting populations through assisted gene flow within the following basins:

- *Thamnophis rufipunctatus* (i) Verde River; (ii) Salt River; (iii) San Francisco River, (iv) Gila River;
- *Thamnophis eques megalops* (i) Bill Williams River; (ii) Verde River; (iii) Salt River, (iv) Santa Cruz River, (v) Gila River.

We provide some site-specific management suggestions and guiding principles for consideration that relate to assisted gene flow for both species in tables 11 and 12. Because most populations are not large, management using reciprocal translocations and multiple sources, where possible, could alleviate the risk of further impacts to source populations. However, in some cases, nearby source populations are unknown or do not exist (for example, Tonto Creek in the Salt River for *T. e. megalops*), so it may be necessary to source individuals from the next most proximate drainage basin, matching recipient and source environments as closely as possible, and choosing less genetically divergent sites and higher diversity sites.

Provided the necessary habitat restoration takes place, efforts to expand population presence within some drainage basins through translocation of snakes into unoccupied areas, either from captive bred colonies or through wild-to-wild translocations of existing populations, could introduce more spatial heterogeneity of gene diversity in each drainage basin. Continued ecological restoration efforts to increase available habitat, reduce predation threats, and increase population sizes are recommended for the long-term success of any genetic rescue program, especially considering climate change (Hedrick and others, 2014; Giermakowski and others, 2015; Frankham and others, 2017). We also suggest that ongoing genetic monitoring of the populations that were evaluated in this study can help evaluate population trends over time. Genetic monitoring of any focal populations where genetic rescue efforts take place will also enable quantifying any genetic changes that may be associated with observable fitness effects (such as increase in population growth and size), assisted gene flow efforts, or further environmental change.

Table 11. Specific sites with low diversity (low effective sizes, bottlenecks, and elevated inbreeding coefficients, F_{IS}) and potential sources for consideration that relate to assisted gene flow in *Thamnophis rufipunctatus*.

Drainage	Site	$N_e < 100$	Population bottleneck	Elevated $F_{IS} > 0.05$ (STACKS/PYRAD)	Potential sources for assisted gene flow
Verde River	Oak	24	Yes	Yes (0.03/0.09)	Multiple sites within the Verde River. Reciprocal translocation of individuals from subsites along Verde River (upper versus lower river populations). Future surveys within the upper Verde River (Prospect Point, Yavapai County, Arizona) may be useful to provide additional sources of genetic variation, either for captive colonies or translocation efforts into Oak Creek.
Salt River	Canyon Creek	19	Yes	Yes (0.00/0.11)	Reciprocal translocation between Canyon Creek and Black River sites could likely be beneficial for long-term population viability at both sites. Future surveys within Tonto Creek and intervening Tribal Lands may provide additional sources of genetic variation.
	Black	42	Yes	Yes (0.02/0.13)	See Canyon Creek (above).
San Francisco River	Blue River	16	Yes	Yes (0.04/0.11)	Multiple sites within the San Francisco River drainage basin: Tularosa River, Whitewater (if still extant), and (or) Saliz Creek.
	Tularosa	30	Yes	Yes (0.04/0.12)	Multiple sites within the San Francisco River drainage basin: Blue River and Whitewater River (if still extant) and (or) Saliz Creek.
Gila River	Middle Fork Gila	13	Yes	Yes (0.05/0.08)	Multiple sites within the Gila River drainage basins: Diamond Creek, West Fork Gila, Saliz Creek (see appendix 1), or some combination thereof.

Table 12. Specific sites with low diversity (low effective sizes, bottlenecks, and elevated inbreeding coefficients, F_{IS}) and potential sources for consideration that relate to assisted gene flow in *Thamnophis eques megalops*.

Drainage	Site	$N_e < 100$	Population bottleneck	Elevated $F_{IS} > 0.05$ (STACKS/PYRAD)	Potential sources for assisted gene flow
Bill Williams River	Centennial Wash	?	No	No (−0.04/−0.01)	Reciprocal translocations among all sites within the Bill Williams River drainage could be used to increase effective sizes, reduce inbreeding potential, and long-term viability of populations.
	Big Sandy	15.8	No	No (0.00/0.02)	See Centennial Wash above.
	Santa Maria	63.4	Yes	No (−0.01/0.01)	See Centennial Wash above.
Verde River	Camp Verde	73.6	Yes	No (0.05/0.05)	Reciprocal translocation among all sites within the Verde River drainage could be used to increase effective sizes, reduce inbreeding potential, and long-term viability of populations.
	Tavasci Marsh	?	Yes	No (0.02/0.02)	See Camp Verde above.
	Dead Horse Ranch	81.5	Yes	No (0.02/0.04)	See Camp Verde above.
	Page Springs	63.5	Yes	No (0.01/0.02)	See Camp Verde above.
Salt River	Tonto Creek	15.1	Yes	Yes (0.07/0.08)	No other populations within the Salt River are currently available as donor sites. Phylogenetic data (RADseq and mitochondrial) indicate Salt River populations are more closely related to the Verde River than to populations within the Santa Cruz River; F_{ST} estimates support this relationship also.
Lower Gila River	Gila River	—	—	No (0.033/−0.009)	Very little information exists about the viability of populations in New Mexico. Recent samples between Cliff and Gila indicate viable populations may still exist. More information is needed before repatriation and (or) genetic translocations strategies can be developed.

Data Accessibility

The RADseq data used for this study are available from the NCBI Sequence Read Archive: BioProject ID PRJNA453585.

References Cited

Alfaro, M.E., and Arnold, S.J., 2001, Molecular systematics and evolution of *Regina* and the Thamnophiine snakes: Molecular Phylogenetics and Evolution, v. 21, p. 408–423.

Axelrod, D.I., 1979, Age and origin of Sonoran Desert vegetation: Occasional Papers of the California Academy of Sciences, v. 132, p. 1–74.

Beugin, M., Gayet, T., Pontier, D., Devillard, S., and Jombart, T., 2018, A fast likelihood solution to the genetic clustering problem: Methods in Ecology and Evolution, v. 9, p. 1006–1016.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A., 2013, Stacks—An analysis tool set for population genomics: Molecular Ecology, v. 22, p. 3124–3140.

Charlesworth, B., 2009, Effective population size and patterns of molecular evolution and variation: Nature, v. 10, p. 195–205.

Cornuet, J.-M., and Luikart, G., 1996, Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data: Genetics, v. 144, p. 2001–2014.

de Queiroz, A., Lawson, R., and Lemos-Espinal, J.A., 2002, Phylogenetic relationships of North American garter snakes (*Thamnophis*) based on four mitochondrial genes—How much DNA sequence is enough?: Molecular Phylogenetics and Evolution, v. 22, p. 315–329.

Do, C., Waples, R.S., Peel, D., Macbeth, G.M, Tillett, B.J., and Ovenden, J.R., 2014, NeEstimator v2–Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data: Molecular Ecology Resources, v. 14, p. 209–214.

Doak, D.F., Himes Boor, G.K., Bakker, V.J., Morris, W.F., Louthan, A. Morrison, S.A., Stanley, A., and Crowder, L.B., 2015, Recommendations for improving recovery criteria under the US Endangered Species Act: Bioscience, v. 65, p. 189–199,

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A., 2012, Bayesian phylogenetics with BEAUti and the BEAST 1.7.: Molecular Biology and Evolution, v. 29, p. 1969–1973.

Eaton, D.A.R., 2014, PYRAD—Assembly of de novo RADseq loci for phylogenetic analysis: Bioinformatics, v. 30, p. 1844–1849.

Edgar, R.C., 2004, MUSCLE—Multiple sequence alignment with high accuracy and high throughput: Nucleic Acids Research, v. 32, p. 1792–1797.

Elgar, M.A., and Clode, D., 2001, Inbreeding and extinction in island populations: A cautionary note: Conservation Biology, v. 15, p. 284–286.

Emmons, I.D., Nowak, E.M., and Lauger, K., 2016, Prey availability and foraging events of the Northern Mexican gartersnake (*Thamnophis eques megalops*) in north-central Arizona: Herpetological Review, v. 47, p. 485–486.

Fischer J., Lindenmayer D.B., 2007, Landscape modification and habitat fragmentation—A synthesis: Global Ecology Biogeography, v. 16, p. 265–280.

Fitzpatrick, S.W., Gernerich, J.C., Angeloni, L.M., Bailey, L.L., Broder, E.D., Torres-Dowdall, J., Handelsman, C.A., López-Sepulcre, A., Reznick, D.N., Ghalambor, C.K., and Funk, W.C., 2016, Gene flow from an adaptively divergent source causes rescue through genetic and demographic factors in two wild populations of Trinidadian guppies: Evolutionary Applications, v. 9, p. 879–891.

Frankham, R., 2005, Genetics and extinction: Biological Conservation, v. 126, p. 131–140.

Frankham, R., 2015, Genetic rescue of small inbred populations—Meta-analysis reveals large and consistent benefits of gene flow: Molecular Ecology, v. 24, p. 2610–2618.

Frankham, R., and Ralls, K., 1998, Conservation biology—Inbreeding leads to extinction: Nature, v. 392, p. 441–442.

Frankham, R., Bradshaw, C.J.A., and Brook, B.W., 2014, Genetics in conservation management—Revised recommendations for 50/500 rules, Red list criteria and population viability analyses: Biological Conservation, v. 170, p. 56–63.

Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R., and Fenster, C.B., 2011, Predicting the probability of outbreeding depression: Conservation Biology, v. 25, p. 465–475.

Frankham, R., Ballou, J.D., Ralls, K., Eldridge, M.D.B, Dudash, M.R., Fenster, C.B., Lacy, R.C., and Sunnucks, P., 2017, Genetic management of fragmented animal and plant populations: Oxford, U.K., Oxford University Press, 432 p.

Giermakowski, J.T., Hatten, J.R., Johnson, M.J., Holmes, J.A., Nowak, E.M., Peters, M., 2015, Final report—Predicting effects of climate change on riparian obligate species in the Southwestern United States: Submitted to U.S. Department of the Interior, Bureau of Reclamation, Desert Landscape Conservation Cooperative, Agreement Number R11AC81532, Albuquerque, N. Mex.: University of New Mexico Digital Repository, http://hdl.handle.net/1928/31796.

Gilpin, M.E., and Soulé, M.E., 1986, Minimum viable populations—Processes of species extinction, chap. 2 of Soulé, M.E., ed., Conservation Biology: Sunderland, Mass., Sinauer Associates Inc, p. 19–34.

Haasl, R.J., and Payseur, B.A., 2011, Multi-locus inference of population structure—A comparison between single nucleotide polymorphisms and microsatellites: Heredity, v. 106, p. 158–171.

Hare, M.P., Nunney, L., Schwartz, M.K., Ruzzante, D.E., Burford, M., Waples, R.S., Ruegg, K., and Palstra, F., 2011, Understanding and estimating effective population size for practical application in marine species management: Conservation Biology, v. 25, p 438–449.

Hartl, D.L., and Clark, A.G., 1989, Principles of population genetics (2d ed.): Sunderland, Mass., Sinauer Associates, xii, 682 p.

Hedrick, P.W., and Fredrickson, R., 2010, Genetic rescue guidelines with examples from Mexican wolves and Florida panthers: Conservation Genetics, v. 11, p. 615–626.

Hedrick, P.W., Peterson, R.O., Vucetich, L.M., Adams, J.R., and Vucetich, J.A, 2014, Genetic rescue in Isle Royale wolves—Genetic analysis and the collapse of the population: Conservation Genetics, v. 15, p. 1111–1121.

Hibbitts, T.J., Painter, C.W., and Holycross, A.T., 2009, Ecology of a population of the Narrow-headed garter snake (Thamnophis rufipunctatus) in New Mexico—Catastrophic decline of a river specialist: Southwestern Naturalist, v. 54, p. 461–467.

Holycross, A.T., Burger, W.P., Nigro, E.J., and Brennan, T.C., 2006, Surveys for Thamnophis eques and Thamnophis rufipunctatus in the Gila River Watershed of Arizona and New Mexico: Final report to Arizona Game and Fish Department.

Jennings, R.D., and Christman, B.L., 2015, Comparisons of the population status of Narrow-headed gartersnakes among sites experiencing post-fire flows and receiving translocations with an unaffected population: Unpublished final report to New Mexico Department of Game and Fish in fulfillment of Share with Wildlife grant.

Johnson, W., Onorato, D., Roelke, M., Land, E.D., Cunningham, M., Belden, R.C., McBride, R., Jansen, D., Lotz, M., Shindle, D., Howard, J., Wildt, D.E., Penfold, L.M., Hostetler, J.A., Oli, M.K., and O'Brien, S.J., 2010, Genetic restoration of the Florida Panther: Science, v. 329, p. 1641–1645.

Jombart, T., Devillard, S., and Balloux, F., 2010, Discriminant analysis of principal components—A new method for the analysis of genetically structured populations: BMC Genetics, v. 11, 94. https://doi.org/10.1186/1471-2156-11-94.

Jombart, T., and Ahmed, I., 2011, Adegenet 1.3-1—New tools for the analysis of genome-wide SNP data: Bioinformatics, v. 27, p. 3070–3071.

Jones, O.R., and Wang, J., 2010, COLONY—A program for parentage and sibship inference from multilocus genotype data: Molecular Ecology Resources, v. 10, p. 551–555.

Kalinowski, S.T., Taper, M.L., and Marshall, T.C., 2007, Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment: Molecular Ecology, v. 16, p. 1099–1106.

Love Stowell, S.M., Pinzone, C.A., and Martin, A.P., 2017, Overcoming barriers to active interventions for genetic diversity: Biodiversity and Conservation, v. 26, p. 1753–1765.

Lowe, W.H., and Allendorf, F.W., 2010, What can genetics tell us about population connectivity?: Molecular Ecology, v. 19, p. 3038–3051.

Luikart, G., and Cornuet, J.-M., 1998, Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data: Conservation Biology, v. 12, p. 228–237

McEachern, M.B., Van Vuren, D.H., Floyd, C.H., May, B., and Eadie, J.M., 2011, Bottlenecks and rescue effects in a fluctuating population of golden-mantled ground squirrels (*Spermophilus laterlis*): Conservation Genetics, v. 12, p. 285–296.

Madsen, T., Ujvari, B., and Olsson, M., 2004, Novels genes continue to enhance population growth in adders (*Vipera berus*): Biological Conservation, v. 120, p. 145–147.

McVay, J.D., Flores-Villela, O., and Carstens, B., 2015, Diversification of North American natricine snakes: Biological Journal of the Linnean Society, v. 116, p. 1–12.

Meffe, G.K., and Vrijenhoek, R.C., 1988, Conservation genetics in the management of desert fishes: Conservation Biology, v. 2, p. 157–169.

Meirmans, P.G., 2015, Seven common mistakes in population genetics and how to avoid them: Molecular Ecology, v. 24, p. 3223–3231.

Metcalfe, S.E., 2006, Late Quaternary environments of the northern deserts and central transvolcanic belt of Mexico: Annals of the Missouri Botanical Garden, v. 93, p. 258–273.

Michalakis, Y., and Excoffier, L., 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci: Genetics, 142, p. 1061–1064.

Miller, J.R., and Hobbs, R.J., 2002, Conservation where people live and work: Conservation Biology, v. 16, p. 330–337.

Miller, J.M., Poissant, J., Hogg, J.T., and Coltman, D.W., 2012, Genetic consequences of genetic rescue in an insular population of bighorn sheep (*Ovis canadensis*): Molecular Ecology, v. 21, p. 1583–1596.

Mills, L.S., and Allendorf, F.W., 1996, The one-migrant-per-generation rule in conservation and management: Conservation Biology, v. 10, p. 1509–1518.

Morrison, R.B., 1991, Quaternary stratigraphic, hydrologic, and climatic history of the Great Basin, with emphasis on Lakes Lahontan, Bonneville, and Tecopa, *in* Morrison, R.B., ed., Quaternary nonglacial geology—Conterminous U.S.: Geological Society of America, v. K-2, p. 283–321.

Nei, M., Maruyama, T., and Chakraborty, R., 1975, The bottleneck effect and genetic variability in populations: Evolution, v. 29, p. 1–10.

Nowak, E.M., 2006, Monitoring surveys and telemetry of narrow-headed gartersnakes (*Thamnophis rufipunctatus*) in Oak Creek, Arizona: Unpublished final report to Arizona Game and Fish Department Heritage Fund, IIPAM Program, U.S. Geological Survey Southwest Biological Science Center Colorado Plateau Research Station, Flagstaff, Ariz.

Nowak, E.M., and Santana-Bendix, M.A., 2002, Status, distribution, and management recommendations for the narrow-headed gartersnake (*Thamnophis rufipunctatus)* in Oak Creek, Arizona: Unpublished final report to Arizona Game and Fish Department in fulfillment of Heritage Grant I99007, U.S. Geological Survey Forest and Rangeland Ecosystem Science Center, Colorado Plateau Field Station, Flagstaff, Ariz.

Nowak, E.M., Drost, C.A., and Mahoney, S.M., 2017, Effects of wildfires on threatened narrow-headed gartersnakes: Unpublished final report to U.S. Geological Survey/U.S. Fish and Wildlife Service SSP-QRP program, U.S. Geological Survey Southwest Biological Science Center, Flagstaff, Ariz.

Paris, J.R., Stevens, J.R., and Catchen, J.M., 2017, Lost in parameter space—A road map for stacks: Methods in Ecology and Evolution, v. 8, p. 1360–1373.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E., 2012, Double Digest RADseq—An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species: PLoS ONE, 7, e37135.

Pierson, J.C., Coates, D.J., Oostermeijer, J.G.B., Beissinger, S.R., Bragg, J.G., Sunnucks, P., Schumaker, N.H., and Young, A.G., 2016, Genetic factors in threatened species recovery plans on three continents: Frontiers in Ecology and the Environment, v. 14, p. 433–440.

Pritchard, J.K., Stephens, M., and Donnelly, P., 2000, Inference of population structure using multilocus genotype data: Genetics, v. 155, p. 945–959.

Ralls, K., Ballou, J.D., Dudash, M.R., Eldridge, M.D.B., Fenster, C.B., Lacy, R.C., Sunnucks, P., and Frankham, R., 2017, Call for a paradigm shift in the genetic management of fragmented populations: Conservation Letters, v. 0, p. 1–6.

Rambaut, A., Drummond, A. J., and Suchard, M., 2014, Tracer v1.6: Molecular evolution, phylogenetics and epidemiology, accessed January 11, 2017, at http://tree.bio.ed.ac.uk/software/tracer.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D., Darling, A., H¨ohna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P., 2011, MrBayes 3.2—Efficient Bayesian phylogenetic inference and model choice across a large model space: Systematic Biology, v. 61, p. 539–542.

Rosen, P.C., 1993, Status of the Narrow-headed Garter Snake (*Thamnophis rufipunctatus*) on Eagle Creek and Blue River, Arizona, *in* Fernandez, P., and Rosen, P.C., Effects of the introduced crayfish *Orconectes virilis* on native aquatic herpetofauna in Arizona: Arizona Game and Fish Department Heritage Report, IIPAM I94054, Phoenix, Ariz., p. 66.

Rosen, P.C., and Schwalbe C.R., 1988, Status of the Mexican and Narrow-headed gartersnakes (*Thamnophis eques megalops* and *Thamnophis rufipunctatus rufipunctatus*) in Arizona: Unpublished report from Arizona Game and Fish Department (Phoenix, Ariz.) to U.S. Fish and Wildlife Service, Albuquerque, N. Mex.

Rossman, D.A., Ford, N.B., and Seigel, R.A., 1996, *Thamnophis rufipunctatus*, Narrow-headed gartersnake, *in* The Gartersnakes: Evolution and ecology: Norman, Okla., University of Oklahoma Press, p. 241–488.

Schönhuth, S., Blum, M.J., Lozano-Vilano, L., Neely, D.A., Varela-Romero, A., Espinos, H., Perdices, A., and Mayden, R., 2011, Inter-basin exchange and repeated headwater capture across the Sierra Madre Occidental inferred from the phylogeography of Mexican stonerollers: Journal of Biogeography, v. 38, p. 1406–1421.

Shafer, A.B., Peart, C.R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C.W., and Wolf, J.B.W., 2017, Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference: Methods in Ecology and Evolution, v. 8, p. 907–917.

Shaffer, M.L., 1981, Minimum population sizes for species: Bioscience, v. 31, p. 131–134.

Smith, M.L., and Miller, R.R., 1986, The evolution of the Rio Grande Basin as inferred from its fish fauna, *in* Hocutt, C.H., and Wiley, E.O., eds., The zoogeography of North American freshwater fishes: New York, Wiley-Interscience Publications, p. 457–485.

Smyser, T.J., Johnson, S.A., Page, L.K., Hudson, C.M., and Rhodes, O.E, 2013, Use of experimental translocations of Allegheny woodrat to decipher causal agents of decline: Conservation Biology, v. 27, p. 752–762.

Tajima, F., 1989, The effect of change in population size on DNA polymorphism: Genetics, v. 123, p. 597–601.

U.S. Fish and Wildlife Service, 2013, Endangered and threatened wildlife and plants; Designation of critical habitat for the Northern Mexican gartersnake and Narrow-headed gartersnake: Federal Register, v. 78, p. 41549–41608.

U.S. Fish and Wildlife Service, 2014, Endangered and threatened wildlife and plants; Threatened status for the Northern Mexican gartersnake and Narrow-headed gartersnake: Federal Register, v. 79, p. 38678–38746.

Waples, R.S., and Do, C., 2008, LDNE—A program for estimating effective population size from data on linkage disequilibrium: Molecular Ecology Resources, v. 8, p. 753–756.

Whiteley, A.R., Fitzpatrick, S.W., Funk, W.C., and Tallmon, D.A., 2015, Genetic rescue to the rescue: Trends in Ecology and Evolution, v. 30, p. 42–49.

Weeks, A.R., Sgrò, C.M., and Young, A.G., 2011, Assessing the benefits and risks of translocations in changing environments—A genetic perspective: Evolutionary Applications, v. 4, p. 709–725.

Weeks, A.R., Heinze, D., Perrin, L., Stoklosa, J., Hoffmann, A.A., van Rooyen, A., Kelly, T., and Mansergh, I., 2017, Genetic rescue increases fitness and aids rapid recovery of an endangered marsupial population: Nature Communications, v. 8, p. 1071–1077.

Weir, B.S., and Cockerham, C.C., 1984, Estimating F-Statistics for the analysis of population structure: Evolution, v. 38, p. 1358.

Wood, D.A., Vandergast, A.G., Lemos-Espinal, J.A., Fisher, R.N., and Holycross, A.T., 2011, Refugial isolation and divergence in the Narrowheaded gartersnake species complex (*Thamnophis rufipunctatus*) as revealed by multilocus DNA sequence data: Molecular Ecology, v. 20, p. 3856–3878.

Wright, S., 1931, Evolution in Mendelian populations: Genetics, v. 16, p. 97–159.

Glossary

Note: Some definitions adapted from National Human Genome Research Institute Talking Glossary of Genetic Terms at https://www.genome.gov/glossary/; Hartl and Clark, 1989; and Frankham and others, 2017.

Adaptive Referring to a genotype or trait conferring high relative fitness to individuals possessing it.

Allele One particular form of a gene (diploid organisms have two of each).

Assisted migration The managed movement of individuals into populations to increase genetic variation and relative fitness.

Base pair A single "position" (or single nucleic acid) on a strand of DNA containing an adenine, cytosine, guanine, or thymine nucleobase. Nuclear DNA is double stranded with complimentary bases or base pairs (guanine-cytosine and adenine-thymine) that allow the DNA helix to maintain a regular helical structure.

Dispersal The movement of individuals from their birth site to their breeding site ("natal dispersal"), as well as the movement from one breeding site to another ("breeding dispersal").

DNA extraction Process of purifying genomic DNA from a tissue or other sample. Most processes include both physical and chemical extraction steps.

Effective population size (N_e) The average number of individuals in a population that contribute genes to succeeding generations. This number generally is lower than the census population size.

Fitness The net effect of viability, mating success, and fecundity that determines number of progeny.

Gene Each gene is a linear segment of a DNA molecule that includes a specific sequence of paired bases that are arranged on chromosomes. Each gene is responsible for a single inherited property, characteristic, or function of the organism.

Gene flow The movement of genes among populations connected by dispersal and migration. The incorporation of "migrant" genes into a population requires both successful immigration and successful interbreeding.

Genetic bottleneck A sharp reduction in the size of a population. Also called a population bottleneck.

Genetic differentiation A process in which two or more populations accumulate independent genetic changes (DNA sequence mutations) or allele frequency differences through time as a result of reduced gene flow or complete reproductive isolation for some period of time. Estimators of genetic differentiation include Wright's fixation index (F_{ST}) and Weir & Cockerham's index θ. Also called genetic divergence.

Genetic diversity The existing genetic variation within a population or species. Various measures are used to estimate genetic diversity such as determining the number of polymorphic sites across a specified region of DNA sequence or determining the number of heterozygous individuals in a population. Some commonly used genetic diversity measures include polymorphic sites(s), heterozygosity (H), and allelic richness (A_r).

Genetic drift The change in allele frequencies over time due to the chance disappearance of particular alleles as individuals die or do not reproduce. Drift is stronger in smaller populations and leads to greater genetic differentiation among populations and lower diversity within them.

Genetic erosion A process in which small populations become inbred and lose genetic diversity, leading to inbreeding depression, reduced adaptation, and elevated extinction risks.

Genetic rescue An increase in population fitness and increase in genetic diversity due to crossing a population previously suffering from inbreeding and (or) low genetic diversity to another population.

Genetic structure The distribution of genotypes within and among populations.

Genotype A genetic profile indicating the particular alleles present at one or more loci within an organism.

Heterozygosity A diploid organism is heterozygous at a gene locus when there are two different alleles of a gene present. Heterozygosity measures the proportion of heterozygous individuals for a locus (or set of loci) in a population.

Inbreeding depression Reduced fitness of offspring as a result of mating of individuals related by descent (brother x sister, father x daughter, cousin x cousin, etc.).

Loci (or markers) Gene regions that are scattered throughout the genome and not physically linked.

Local adaptation The process by which organisms in a particular locality have differentially evolved as compared to other localities in response to selective pressures imposed by some biotic or abiotic aspect of their local environment.

Migration The physical movement of individuals from one area to another; can be seasonal.

Next generation sequencing (or High throughput sequencing) Sequencing is the process of determining the precise order of nucleotides within a DNA molecule. High-throughput sequencing can sequence very large quantities of DNA at one time. Platforms include Illumina® HiSeq, MiSeq (Illumina, Inc.), and PacBio® RS (Pacific Biosciences of California, Inc.), among others.

Nucleotide diversity The average proportion of nucleotides that differ between any randomly sampled pair of sequences.

Outbreeding depression Reduced fitness of offspring as a result of matings between genetically divergent individuals.

Phylogenetic lineage (clade) A branch on a phylogenetic tree that consists of a common ancestor (node of the branch) and all of its descendants (terminal taxa at the tips of the branches). Clades are nested, as each branch separates into smaller branches. These splits reflect the evolutionary history of clade divergence and independent evolution.

Population A group of related individuals in the same geographic area that freely interbreed.

Population genetic structure The accumulated genetic differences between groups of individuals that do not freely interbreed. Physical barriers to migration and limited dispersal can cause genetic structuring of populations. The amount of genetic structure is inversely related to the amount of gene flow.

Private allele An allele found in only a single population.

Restriction enzyme An enzyme that cuts DNA molecules at a specific sequence.

Selection Differential survival and reproduction among individuals due to heritable trait differences.

Sequencing The process of determining the precise order of nucleotides within a DNA molecule.

SNP—Single Nucleotide Polymorphism A single base mutation at a specific locus usually consisting of two alleles. In a diploid organism, an individual can be either homozygous for one or another allele copy, or heterozygous (containing one of each allele copy). Many SNPs can be combined to provide an individual genetic fingerprint or genotype.

Appendix 1. Determining the Origin of Released Snakes into Saliz Creek, New Mexico

In 2012, 18 individuals of *T. rufipunctatus* were released into Saliz Creek, a tributary of the San Francisco River, New Mexico. This site was believed to be devoid of any natural population on the basis of three surveys conducted by Lyndsey Helikson. However, the exact origin of the released snakes is not known, but they are believed to have been salvaged from either the Middle Fork Gila, Whitewater Creek, or some combination of the two sites. Subsequent surveys at Saliz Creek by Bruce Christman and Erika Nowak produced a total of three individuals in 2013, two snakes in 2015, and two snakes 2016, and tissue samples were taken of each of these snakes. We used the SNP dataset generated in this study to determine the possible origin of the released snakes using phylogenetic and Bayesian clustering analyses (STRUCTURE; Pritchard and others, 2000).

Bayesian analysis in STRUCTURE (fig. 1–1) revealed that individuals from Saliz Creek have admixed ancestry coefficients between Gila River (Middle Fork Gila) and San Francisco River (Whitewater Creek).
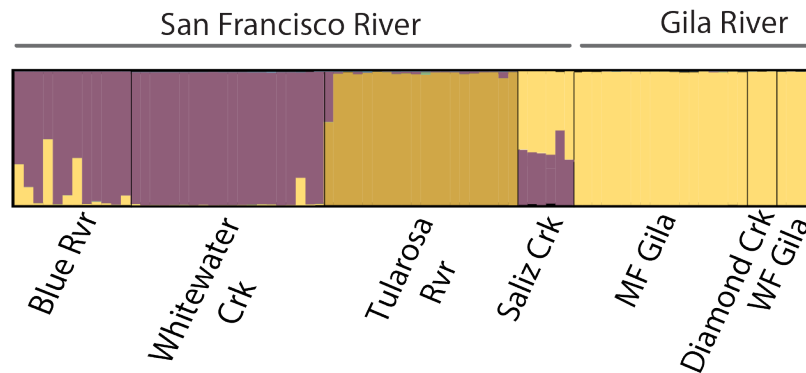
Figure 1–1. Cluster assignments of sampling sites for *Thamnophis rufipunctatus* to determine the source of Saliz Creek individuals.

Similarly, phylogenetic relationships inferred from Bayesian analysis showed Saliz Creek individuals nested within a clade containing individuals from Middle Fork Gila and other samples from West Fork Gila, and a single individual that nested within the clade containing individuals from Whitewater Creek (fig. 1–2), indicating that both Middle Fork Gila and Whitewater Creek individuals were the probable source of the translocations.
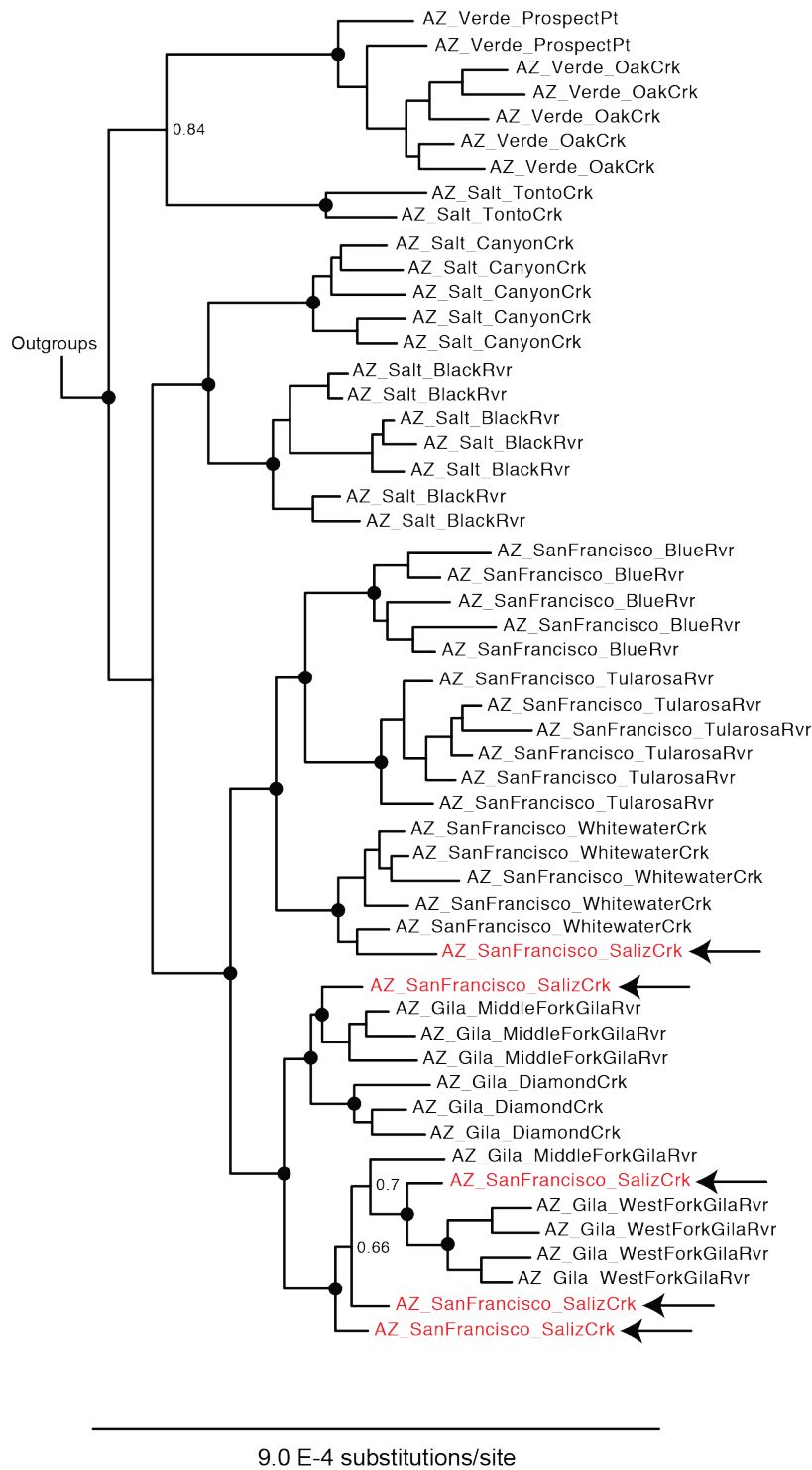
Figure 1–2. Phylogenetic tree of *Thamnophis rufipunctatus* sampling sites using MrBayes, including samples collected from Saliz Creek (red colored font, black arrows). *Thamnophis unilabialis* and *T. nigronuchalis* clades were trimmed out of the tree for clarity. Black dots indicate branch support of ≥ 0.95 posterior probability.

Appendix 2. Summary Statistics Comparisons between STACKS and PYRAD Pipeline Methods

Shafer and others (2017) showed that processing of RADseq data using different methods can potentially have a large effect on resulting genetic diversity summary statistics. On the basis of their recommendations, we evaluated the robustness of our summary statistics for both species of gartersnakes using datasets generated from the PYRAD and STACKS bioinformatics pipelines. For the PYRAD pipeline, we generated clusters of homologous loci using a clustering threshold of 90 percent and allowing up to 30 percent missing data. We discarded consensus sequences with more than 20 percent heterozygous sites and more than two alleles for an individual. We exported the PYRAD dataset in a variable call format (VCF) file. Once imported into STACKS, we applied filtering thresholds similar to those used to generate the stacks dataset (table 2–1). For each dataset, we calculated the following summary statistics for comparisons: number of private alleles (P), mean observed heterozygosity (H_{obs}), mean expected heterozygosity (H_{exp}), and the inbreeding coefficient (F_{IS}).

Table 2–1. Summary of RADseq datasets assembled for *Thamnophis rufipunctatus* and *T. eques megalops* using the vcf pyrad and stacks pipelines.

[-p, minimum number of populations a locus must be present in to process a locus; -r, minimum percentage of individuals in a population required to process a locus in a populations; min_maf, minimum minor allele frequency required to process a nucleotide site at a locus; Loci, total number of homologous loci; SNP, number of single nucleotide polymorphisms recovered from each locus (using a single SNP from each locus)]

Species	Dataset	-p	-r	min_maf	Loci	SNPs
Thamnophis rufipunctatus	PYRAD	7	80	0.05	821	85
	STACKS	7	80	0.05	2625	253
Thamnophis eques megalops	PYRAD	7	80	0.05	2082	207
	STACKS	7	80	0.05	2307	537

We found that STACKS recovered more loci and SNPs than datasets generated in PYRAD (table 2–1). Similarly, estimates of diversity were generally higher for datasets generated using the STACKS pipeline (figs. 2–1 and 2–2). For *T. rufipunctatus*, differences in summary statistics among datasets were generally proportional among pipelines despite PYRAD returning lower estimates. However, we observed non-proportional differences between pipelines for *T. e. megalops*, where heterozygosity estimates were much higher using the stacks pipeline (fig. 2–2). We also found that inbreeding coefficients were generally higher in datasets generated using the PYRAD pipeline, but standard errors for this measure were high for both pipelines.
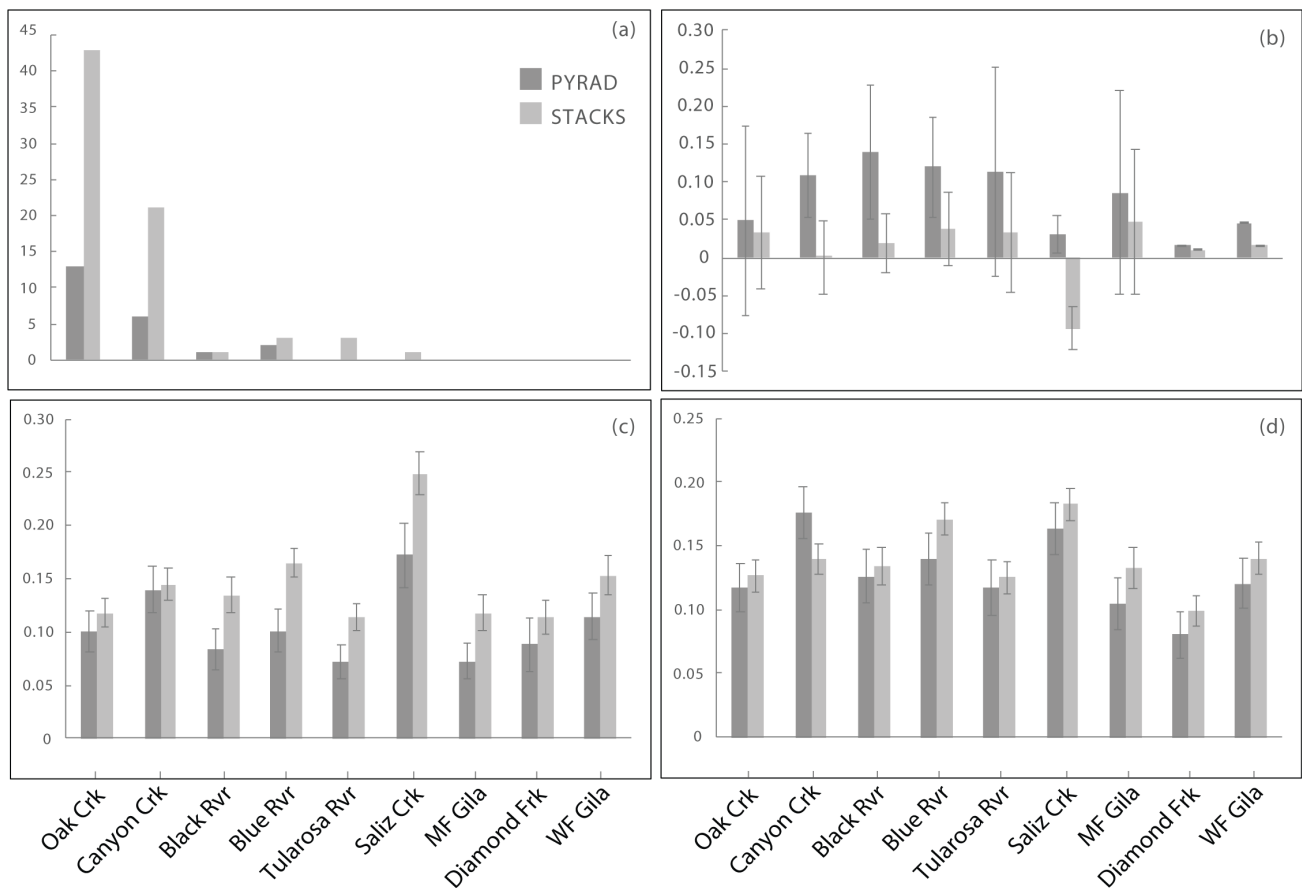
Figure 2–1. Summary statistics comparisons between PYRAD and STACKS pipelines at sampling sites across the lower Colorado River Basin for *Thamnophis rufipunctatus* with standard errors shown around each mean: (a) number of private alleles, *P*; (b) inbreeding coefficients, F_{IS}; (c) observed heterozygosity, H_{obs}; and (d) expected heterozygosity, H_{exp}.
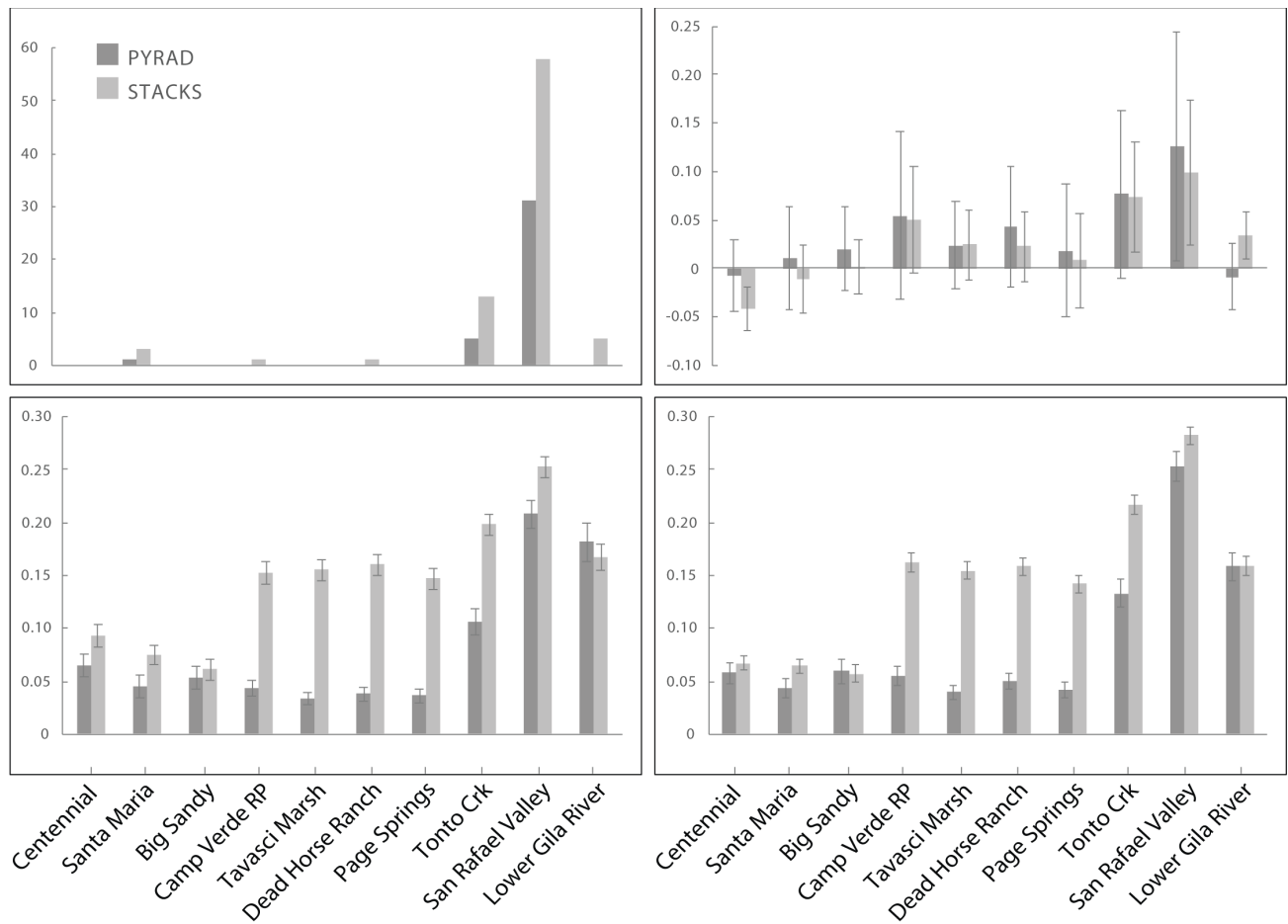
Figure 2–2. Summary statistics comparisons between PYRAD and STACKS pipelines at sampling sites across the lower Colorado River Basin for *Thamnophis eques megalops* with standard errors shown around each mean: (a) number of private alleles, *P*; (b) inbreeding coefficients, F_{IS}; (c) observed heterozygosity, H_{obs}; and (d) expected heterozygosity, H_{exp}.

Appendix 3. Parentage and Sibship Relationships among Captive Snakes

In 2014, the Phoenix Zoo paired a captive born *Thamnophis rufipunctatus* female (Rhiannon) with two unrelated *T. rufipunctatus* males (Manzanita and Jake). This pairing resulted in 20 neonates; however, the exact sire was not known. According to observations, Phoenix Zoo staff thought the likely sire was Jake, but they could not rule out Manzanita. We used the SNP data and a likelihood-based approach to assign parentage to the neonates using the program Cervus v3.0.7 (Kalinowski and others, 2007). Cervus calculates the log-likelihood of each candidate parent being the true parent relative to an arbitrary individual and then calculates the difference between the two most likely parents (LOD score), a statistical confidence is assigned using simulation analysis. We evaluated parentage assignment for 12 of the 20 siblings from Rhiannon from the Phoenix Zoo using two datasets generated in the PYRAD pipeline: (1) a "full" dataset containing 215 SNPs and (2) a "reduced" dataset containing 79 SNPs. To generate the datasets, we used different cutoff thresholds of the polymorphic information content (PIC) score, which was generated by Cervus for each locus. For the "full" dataset we used a PIC > 0.305, and for the "reduced" dataset we used a PIC > 0.369. For each dataset, allele frequencies were generated from the real data, and simulated datasets were generated with parameters set at 100,000 offspring, with 100 percent of candidate parents sampled and a total proportion of SNPs typed over all individuals of 0.874 (full) and 0.884 (reduced), mistyping error rates = 0.01 and a minimum number of SNPs typed of 108 (full) and 40 (reduced). We set the confidence levels at 95 percent (Strict) and 80 percent (Relaxed).

We also generated a SNP dataset to estimate paternity for a captive group of *T. rufipunctatus* siblings that were bred at Northern Arizona University (NAU). Northern Arizona University paired a female (11346) with two male siblings (11341, 11342), and the following year she gave birth to six neonates. We used Cervus to determine the paternity of the six neonates and whether multiple paternity was possible. We used the same Cervus parameters as above except that the total proportion of loci typed over all individuals of 0.985 (84 SNPs), and we implemented the inbreeding option for the simulated data to account for relatedness between parents and offspring. We set the confidence levels at 95 percent (strict) and 80 percent (relaxed).

For *Thamnophis eques megalops*, we obtained tissue samples for 15 neonates that were born from a wild-caught female at Dead Horse Ranch State Park (DHRSP). We generated a SNP dataset within PYRAD (151 SNPs) to estimate sibship relationships among the 15 neonates and determine whether these snakes were from a single sire or multiple sires. We used the program Colony (Jones and Wang, 2010) rather than Cervus because we did not have tissue samples from the possible sires. For each Colony run, we specified a mating system allowing for male polygamy and female monogamy without inbreeding, and used the Full-likelihood method and no sibship prior. Analyses were repeated with different run lengths to confirm the recovered results.

Thamnophis rufipunctatus

Using the SNP datasets generated in Cervus, we were able to assign Manzanita as the most likely father of the neonates that were born at the Phoenix Zoo. Paternity was assigned with higher confidence using the reduced dataset, but both datasets resulted in assigning Manzanita (PHX12303) as the father with 95 percent confidence (table 3–1). There was also no evidence for multiple paternity from this sample. Results for the paternity analysis for the colony housed at NAU are reported in table 3–2. We found no evidence for multiple paternity in this colony, and all offspring were paired with PHX11342 as the most likely candidate father with 95-percent confidence.

Table 3–1. Cervus parentage assignments for the Phoenix Zoo Narrow-headed gartersnakes using full and reduced SNP datasets.

[For each offspring, the mother (Rhiannon: PHX11345) was included in the analysis, and the total number of mismatching loci and LOD score are reported for the trio (offspring, mother, and the candidate father); Candidate fathers were Manzanita: PHX12303 and Jake: PHX10809; * = 95% confidence in paternal assignment]

Full dataset (215 loci)				Reduced dataset (79 loci)			
Offspring	Candidate father	Trio loci mismatch	Trio LOD score	Offspring	Candidate father	Trio loci mismatch	Trio LOD score
PHX12627	PHX12303*	5	–9.37	PHX12627	PHX12303*	1	–1.71
	PHX10809	14	–42.50		PHX10809	4	–10.80
PHX12630	PHX12303*	2	2.96	PHX12630	PHX12303*	1	3.17
	PHX10809	16	–54.00		PHX10809	5	–13.90
PHX12631	PHX12303*	3	–7.09	PHX12631	PHX12303*	0	1.99
	PHX10809	16	–52.10		PHX10809	2	–6.81
PHX12633	PHX12303*	4	–7.22	PHX12633	PHX12303*	0	2.88
	PHX10809	13	–36.40		PHX10809	5	–17.60
PHX12634	PHX12303*	6	–19.40	PHX12634	PHX12303*	1	–2.42
	PHX10809	17	–50.20		PHX10809	4	–9.22
PHX12636	PHX12303*	3	–4.79	PHX12636	PHX12303*	0	1.74
	PHX10809	13	–36.70		PHX10809	2	–6.61
PHX12638	PHX12303*	6	–16.00	PHX12638	PHX12303*	1	–2.35
	PHX10809	11	–29.60		PHX10809	3	–8.98
PHX12639	PHX12303*	7	–18.50	PHX12639	PHX12303*	1	–2.67
	PHX10809	13	–39.00		PHX10809	2	–4.25
PHX12640	PHX12303*	2	2.52	PHX12640	PHX12303*	0	3.09
	PHX10809	12	–31.80		PHX10809	2	–4.50
PHX12641	PHX12303*	7	–23.50	PHX12641	PHX12303*	1	–3.35
	PHX10809	14	–45.40		PHX10809	3	–8.58
PHX12642	PHX12303*	6	–12.60	PHX12642	PHX12303*	1	–3.75
	PHX10809	13	–33.90		PHX10809	4	–14.00
PHX12643	PHX12303*	3	–2.57	PHX12643	PHX12303*	0	1.43
	PHX10809	22	–79.70		PHX10809	7	–26.80

Table 3–2. Cervus parentage assignments for the NAU Narrow-headed gartersnake captive breeding.

[For each offspring, the mother (PHX113<u>46</u>) was included in the analysis, and the total number of mismatching loci and LOD score are reported for the trio (offspring, mother, and the candidate father); Candidate fathers were siblings of PHX113<u>46</u>; Bold = 95% confidence in paternal assignment]

Offspring	Candidate father	Trio loci mismatch	Trio LOD score
DAW15_176	**PHX11342**	**4**	**–3.55**
	PHX11340	4	–6.06
	PHX11341	7	–16.63
DAW15_177	**PHX11342**	**2**	**4.12**
	PHX11340	4	–7.56
	PHX11341	8	–24.12
DAW15_178	**PHX11342**	**3**	**1.33**
	PHX11340	6	–14.55
	PHX11341	8	–18.95
DAW15_180	**PHX11342**	**3**	**–0.72**
	PHX11340	6	–14.24
	PHX11341	9	–27.29

Thamnophis eques megalops

For Mexican gartersnakes, our Colony analyses inferred 105 full sibling pairwise relationships among the 15 neonates that were born from a wild-caught female at Dead Horse Ranch State Park (DHRSP) with 100 percent confidence (fig. 3–1), indicating a single sire family structure for these neonates. We performed three independent analyses, and each run gave the same result.
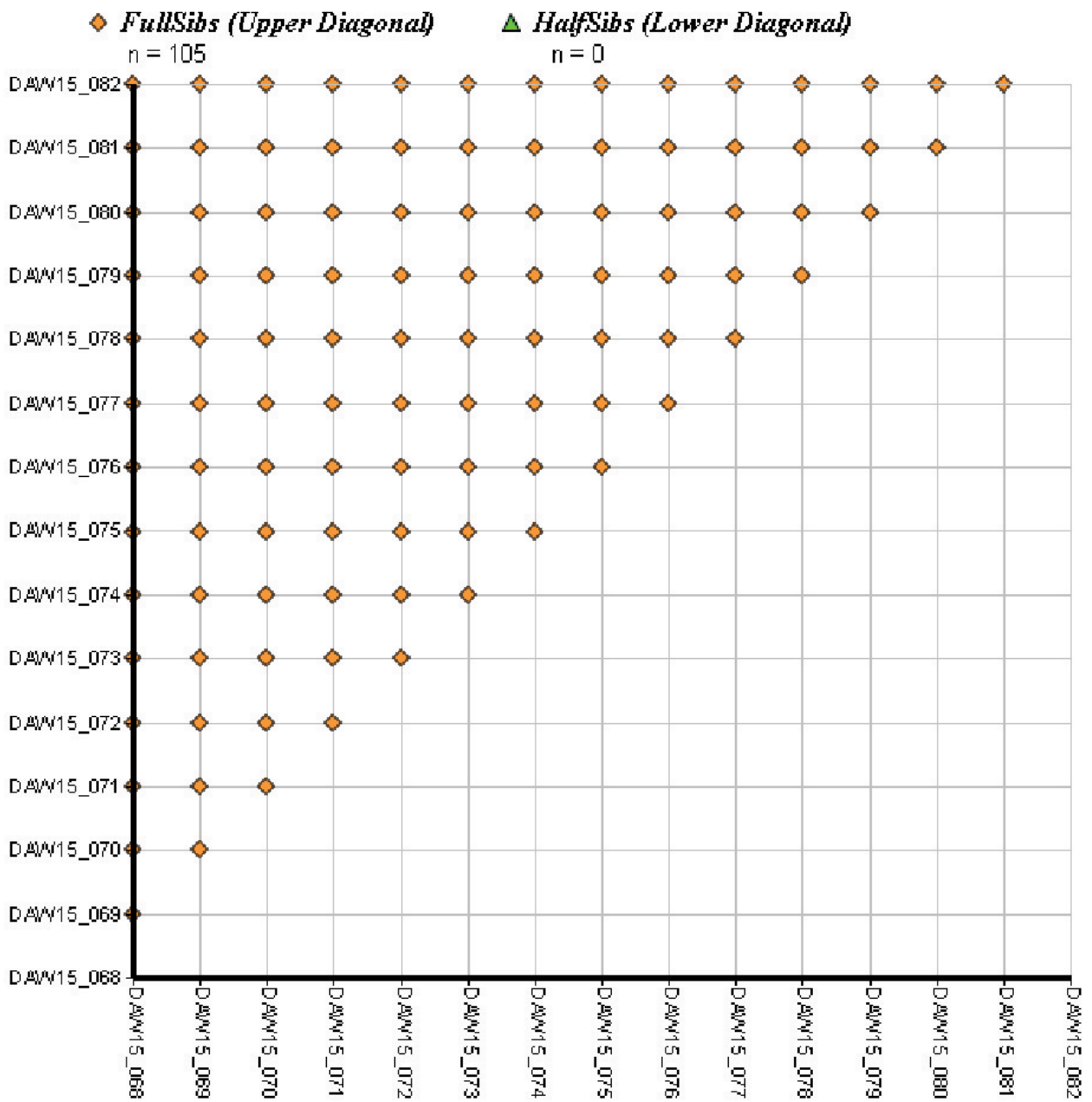
Figure 3–1. Colony analysis inferred 105 full sibling pairwise relationships among the 15 neonates.

USGS

Wood and others—**Conservation Genomics of the Mogollon Narrow-Headed Gartersnake (*Thamnophis rufipunctatus*) and Northern Mexican Gartersnake (*Thamnophis eques megalops*)**—Open-file Report 2018–1141