

Refugial isolation and divergence in the Narrowheaded Gartersnake species complex (*Thamnophis rufipunctatus*) as revealed by multilocus DNA sequence data

DUSTIN A. WOOD,* A. G. VANDERGAST,* J. A. LEMOS ESPINAL,† R. N. FISHER* and A. T. HOLYCROSS‡§

*U.S. Geological Survey, Western Ecological Research Center, San Diego Field Station, 4165 Spruance Road, Suite 200, San Diego, CA 92101, USA, †Laboratorio de Ecología, UBIPRO, Facultad de Estudios Superiores Iztacala, UNAM, A.P. 314, Avenida de los Barrios, No. 1, Los Reyes Iztacala, Tlalnepantla, Estado de México 54090, México, ‡Mesa Community College, 7110 East McKellips Road, Mesa, AZ 85207, USA, §School of Life Sciences, Arizona State University, College and University Streets, Tempe, AZ 85287, USA

Abstract

Glacial–interglacial cycles of the Pleistocene are hypothesized as one of the foremost contributors to biological diversification. This is especially true for cold-adapted montane species, where range shifts have had a pronounced effect on population-level divergence. Gartersnakes of the *Thamnophis rufipunctatus* species complex are restricted to cold headwater streams in the highlands of the Sierra Madre Occidental and southwestern USA. We used coalescent and multilocus phylogenetic approaches to test whether genetic diversification of this montane-restricted species complex is consistent with two prevailing models of range fluctuation for species affected by Pleistocene climate changes. Our concatenated nuDNA and multilocus species analyses recovered evidence for the persistence of multiple lineages that are restricted geographically, despite a mtDNA signature consistent with either more recent connectivity (and introgression) or recent expansion (and incomplete lineage sorting). Divergence times estimated using a relaxed molecular clock and fossil calibrations fall within the Late Pleistocene, and zero gene flow scenarios among current geographically isolated lineages could not be rejected. These results suggest that increased climate shifts in the Late Pleistocene have driven diversification and current range retraction patterns and that the differences between markers reflect the stochasticity of gene lineages (i.e. ancestral polymorphism) rather than gene flow and introgression. These results have important implications for the conservation of *T. rufipunctatus* (sensu novo), which is restricted to two drainage systems in the southwestern US and has undergone a recent and dramatic decline.

Keywords: incomplete lineage sorting, mtDNA, nuDNA, phylogeography, Pleistocene, Sierra Madre Occidental

Received 21 July 2010; revision received 9 June 2011; accepted 14 June 2011

Introduction

The influence of Pleistocene glaciations on species distributions and patterns of divergence in North America

has been the subject of intense interest (Hewitt 1996, 2000). Repeated glacial–interglacial cycling strongly influenced the Pleistocene climate, and these climatic fluctuations have drastically altered species distributions and genetic diversification (Avice 2000; Hewitt 2000). The general pattern for North American temperate taxa indicates displacement and fragmentation of

Correspondence: Dustin A. Wood, Fax: 619 225 6436; E-mail: dawood@usgs.gov

species distributions into more southern glacial refugia because of advancing ice sheets, followed by range expansions northward during warmer interglacial periods as ice sheets contracted. Consequently, the contemporary distributions of many species at northerly latitudes in North America are a result of colonization from southern refugia over the past 10–20 000 years. Challenging this model, patterns from recent investigations across the intermontane west of North America have varied considerably between species, presumably due to the heterogeneity of landscape features (enabling persistence of populations across multiple refugia) and species-specific characteristics (e.g. cold-adapted vs. warm-adapted species), indicating that glacial climatic conditions were not inherently unfavourable or restrictive for all species (Knowles 2000; DeChaine & Martin 2004; Galbreath *et al.* 2009). This has led to two prevailing models of range fluctuation that contrast specific predictions regarding the genetic consequences of glacial–interglacial cycling—the ‘latitude shift’ and ‘archipelago’ models.

These two models of range fluctuation are expected to result in different genetic outcomes. Under the latitude shift model, genetic diversity is expected to decrease with increasing latitude, consistent with dispersal into and out of southern refugia (Hewitt 2004). Demographic estimates and patterns of haplotype variation should also exhibit signatures of population expansion and long-distance dispersal, reflecting postglacial recolonization into northern latitudes as glaciers contracted. In contrast, the archipelago model predicts that genetic diversity should not be strongly correlated with latitude. Rather, population divergence would be promoted mainly by genetic drift because of restricted gene flow across multiple high-elevation refugia. The demographic signatures of the archipelago model should yield population expansion or stability as cool-adapted habitats shifted to lower elevations during glacial maxima, followed by population contractions during warmer interglacial periods as preferred habitats shifted to higher elevations.

Although continental ice sheets and mountain and valley glaciers did not penetrate the lower latitudes of southwestern North America (Porter *et al.* 1983), climatic changes during Pleistocene glaciations undoubtedly affected species in these areas, especially cool-adapted species currently inhabiting high-elevation biomes. The Mexican highlands and mountain isolates (sky islands) of the southwestern USA present an excellent opportunity for examining the genetic consequences of Pleistocene climatic changes on species distribution and diversification. Surprisingly, few studies have focused attention across these southern montane regions (especially the Sierra Madre Occidental) where cold-adapted

species may have responded uniquely to past climatic events, and little is known about the historical biogeography of the Mexican highland species (Smith & Farrell 2005; McCormack *et al.* 2008a, 2010; Bryson *et al.* 2011; Gugger *et al.* 2011). The region is topographically complex, with high north–south oriented mountain complexes separated by deep drainages and low-elevation intermountain basins. This extensive latitudinal and elevational range may have allowed for ephemerally contiguous highland biotas to persist in multiple areas during past climatic shifts and therefore presents an excellent landscape for evaluating the prevailing models of range fluctuation for cool-adapted species.

Across the Mexican highlands and mountain isolates of the southwestern USA, several biomes were strongly influenced by repeated glacial–interglacial cycling. Palynological records and more recent data derived from packrat middens detail a history in which elements of ecological zones shifted dramatically both in latitude and in elevation in response to the climatic transitions at the end of the most recent glacial period (Van Devender 1990). This has resulted in the vividly contrasting habitats across this region observed today at high (temperate montane forests) and low (desert scrub) elevations and likely reflects how cold-adapted species persisted in these southerly distributed mountains during earlier interglacials throughout the Pleistocene (Pielou 1991). Because populations of cold-adapted species in this region likely experienced repeated cycles of down-slope range expansion during cooler glacial periods and up-slope contraction during the warmer interglacials, analysis of the temporal and spatial genetic signatures of connectivity and isolation should exhibit patterns most consistent with the archipelago model of range fluctuation.

Little is known about the phylogeography and genetic structure of the narrowheaded gartersnake (*Thamnophis rufipunctatus*) species complex. These snakes occur primarily in clear, cold-water streams in the Mexican highland of the Sierra Madre Occidental from Durango northward into northern Chihuahua and eastern Sonora, and in the United States, along the Mogollon Rim of Arizona and New Mexico (Thompson 1957; Tanner 1990). Across this range, present-day populations occur in three disjunct isolates (Fig. 1). While a moderate degree of variation in cranial and external measurements, scalation, and color pattern has been described across the three isolates, taxonomic uncertainty remains about the number of species or subspecies that might be recognized in this garter snake species complex (Thompson 1957; Tanner 1985; Rossman 1995). The most recent morphological study recognized two species, *T. rufipunctatus* and *Thamnophis nigronuchalis*; however, previous arrangements have

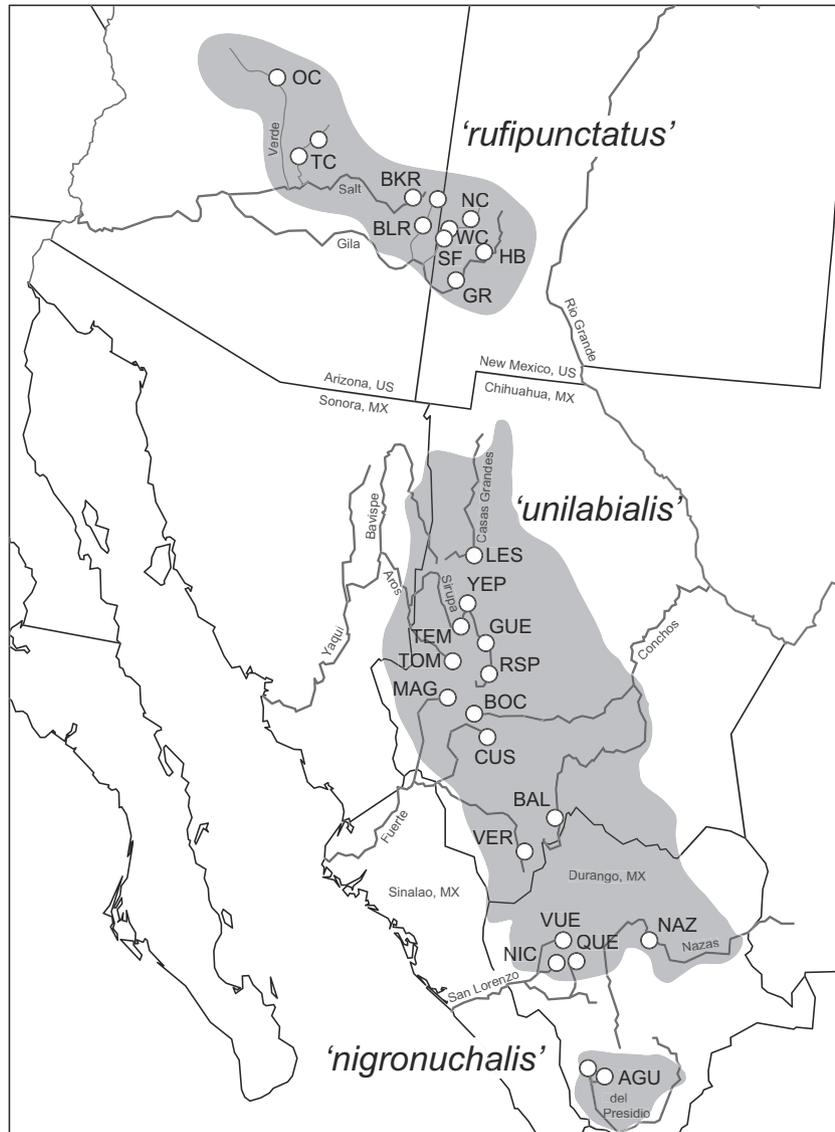


Fig. 1 Distribution of the *Thamnophis rufipunctatus* species complex showing sampled localities across the three geographic isolates. Detailed information for abbreviated localities is found in Appendix I.

classified the three geographic isolates as distinct subspecies under *T. rufipunctatus* (*T. r. rufipunctatus*, *T. r. unilabialis* and *T. r. nigronuchalis*, Tanner 1985; Fig. 1). These snakes are aquatic specialists that feed almost exclusively on fish, use specific habitats for foraging (clear, rock-boulder strewn streams), have evolved several mechanisms for increased underwater visual and foraging capabilities, and display very low dispersal capacity (Flehart 1967; Schaeffel & de Queiroz 1990; Alfaro 2002; Hibbitts & Fitzgerald 2005). Given these specializations and life history characteristics, the *T. rufipunctatus* complex is likely to be highly susceptible to environmental change, especially climate change, providing a model species group for uncovering phylogeographic patterns in this region.

Identifying phylogeographic breaks and the distributions of clades is not only important for understanding the effects of climate change on diversification, but also for conservation management. Populations of *T. rufipunctatus* in the United States are separated from populations in the highlands of Mexico by *c.* 350 km and occur in two primary watersheds (Gila and Salt rivers systems) along the Mogollon Rim in New Mexico and Arizona. These northern populations have experienced severe declines over the past 30 years. Historical surveys indicate that *T. rufipunctatus* has been extirpated from more than 60% of its historic range in the United States, with remaining populations becoming increasingly more isolated and/or experiencing local extirpation (Rosen & Schwalbe 1988; Holycross *et al.*

2006; Hibbitts *et al.* 2009). Given that little is known about the genetic structure of these populations, phylogeographic analysis can provide a measure of genetic diversity and differentiation among these geographic isolates, give a temporal measure of past interchange between populations, and assist in revealing populations that are in need of conservation status assessment.

In this study, we examine multilocus genetic variation and infer the historical biogeography in the *T. rufipunctatus* species complex to explore the effects of Pleistocene climatic changes on phyletic diversification of these montane-restricted species. First we use phylogenetic analyses of DNA sequence data, from both mtDNA and nuDNA loci, to assess whether the three geographic isolates across the species' range each comprise divergent phylogenetic clades consistent with the predictions of the archipelago model (i.e. multiple high-elevation refugia). For recently diverged species, lineage sorting can be incomplete and dispersed throughout the genome, but recent advances in coalescent methods are now available to help detect signals of divergence even before lineages have become reciprocally monophyletic (Knowles & Maddison 2002; Hey & Nielsen 2004). Therefore, we conducted separate analyses of the mtDNA and nuDNA loci to test for incongruence among gene trees and then used several multilocus coalescent methods to jointly estimate divergence times and other demographic quantities to help identify processes responsible for generating the recovered gene tree patterns. Divergence time estimates were used to assess whether clade diversification was consistent with predictions of climatic/environmental change induced by Pleistocene glacial cycles or older events. Given that the preferred habitat of the *T. rufipunctatus* complex likely expanded during glacial episodes and retracted as climate warmed, we used historical demographic estimates of migration among clades (along with estimates of divergence time) to test whether shared haplotypes between geographic isolates resulted from postisolation gene flow or retention of ancestral poly-

morphism. Finally, we examined patterns of population genetic diversity within and among watersheds to whether diversity is correlated with latitude and to provide information for conservation assessment and management priorities.

Materials and methods

Tissue sampling, laboratory techniques and molecular data

We obtained tissue samples from 91 *Thamnophis rufipunctatus* specimens from 25 collection sites (avg. 3.7 individuals per site, range 1–11). These sites encompass all species and/or subspecies ascribed to the *T. rufipunctatus* complex (Fig. 1; 42 *rufipunctatus*; 40 *unilabialis*; 9 *nigronuchalis*). Specimen locality data and voucher information are provided in Appendix I.

We extracted genomic DNA from ethanol-preserved tissues (liver, muscle, blood and tail tips) using standard DNA extraction techniques (Hillis *et al.* 1996). Four genomic fragments (~2817 bp total) representing one mitochondrial protein-coding gene (mtDNA) and three nuclear gene regions were amplified using PCR and sequenced for most individuals. The mitochondrial DNA fragment (1098 bp) contained the entire NADH subunit 1 protein coding gene (ND1) and a short portion of the 16S rRNA coding gene. The three nuclear loci used in the study consisted of an intron fragment flanking the TATA Box (TATA, 753 bp), a short portion of the Vimentin intron 5 (VIM5, 393 bp) and R35 coding gene (573 bp). The rationale for using the three independent nuclear loci was to provide a more rigorous assessment of phylogenetic relationships and help identify instances of incongruence, as single-marker phylogenetic analyses of closely related taxa can be misled by incomplete lineage sorting and/or introgression (Ballard & Whitlock 2004). Primers used for PCR and sequencing are provided in Table 1. Approximately 50–100 ng of total DNA was used as template for PCR

Table 1 Mitochondrial and nuclear primers used in this study

| Primer | Sequence (5'-3') | Locus | Length (bp) | Source |
|-----------|------------------------------|-------------------|-------------|-------------------------------------|
| 16DR | CTACGTGATCTGAGTTCAGACCGGAG | NADH1 | 1098 | Wood <i>et al.</i> 2008 |
| ILE-R | TCTCRGGCACAYTTCATTGTGGT | NADH1 | 1098 | Wood <i>et al.</i> 2008 |
| VIM5F_2 | CAACCAGCCAAGCCAGTC | Vimenton Intron 5 | 393 | Modified from Pyron & Burbrink 2009 |
| VIM6R_2 | GGCGAGCCATTCCTCTT | Vimenton Intron 5 | 393 | Modified from Pyron & Burbrink 2009 |
| TBP-5-6-F | TGTGATGTMAAATCCCTATCMGACTTGA | TATA box Intron 3 | 753 | D. Leavitt personal communication |
| TBP-5-6-R | ACAATTCTGGTTTGATCATTCTGTA | TATA box Intron 3 | 753 | D. Leavitt personal communication |
| R35F | GACTGTGGAYGAYCTGATCAGTGTGG | R35 | 573 | Leaché 2009 |
| R35R | GTAGTGATCCAAGTGACAGTA | R35 | 573 | Leaché 2009 |

in a final volume of 12.5 μ L containing 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.2 μ M of each primer, 1 \times PCR buffer (Invitrogen) and 1 unit of DNA TAQ Polymerase. Thermal cycling conditions consisted of 95 °C for 5 min, 35 cycles of 94 °C for 45 s, 51 °C (ND1), 58 °C (R35) or 60 °C (TATA and VIM5) for 45 s, 72 °C for 75 s, followed by a 72 °C final extension for 8 min. We purified PCR products and directly sequenced on an ABI 3100 capillary system. Sequences were edited using Sequencher™ 4.6 and aligned by eye.

To identify allelic variants within heterozygous individuals, the phase of nuclear genotypes were estimated using a Bayesian approach implemented with the program PHASE 2.1.1 (Stephens *et al.* 2001), accepting results with a probability >90%. Individuals with a probability <90% were cloned using TOPO TA cloning kits (Invitrogen, Carlsbad, CA, USA), with up to five clones sequenced each to assess the gametic phase of the haplotypes. To satisfy the model assumptions across the various methods employed, we tested for recombination in the nuclear loci using the difference of sums of squares (DSS) test in TOPALI version 2.5 (Milne *et al.* 2004). Deviations from neutrality for all loci were tested with the HKA test (Hudson *et al.* 1987) implemented in the program DnaSP version 5 (Librado & Rozas 2009).

Genetic diversity analyses

We used ARLEQUIN v3.1 (Excoffier *et al.* 2005) to assess the genetic diversity of populations within the three geographic isolates and major drainages where sample sizes of five or more were obtained. Genetic diversity was estimated for both mtDNA and nuDNA using a variety of statistics, including number of haplotypes per drainage, haplotype/allelic diversity (Hd), nucleotide diversity (π) and number of polymorphic sites (s). For

both mtDNA and nuDNA data sets, linear regressions of nucleotide diversity (π) for each major drainage vs. latitude were performed to test whether diversity decreased with increasing latitude as would be expected under a model of rapid expansion from southern refugia, or whether diversity has no relation with latitude (the major drainages tested are listed in Table 2). We also estimated pairwise Φ_{CT} and performed an analysis of molecular variance (AMOVA) among geographic isolates to explore whether there existed significant genetic variation among different groupings. The groupings analysed were a two species hypothesis (*T. rufipunctatus* vs. *T. nigronuchalis*) and a three species hypothesis equivalent to each geographic isolate.

Gene tree estimation

Bayesian phylogenetic analyses for mtDNA and nuDNA gene tree analyses were conducted using the program MRBAYES version 3.1.2 (Huelsenbeck & Ronquist 2001). We chose to use MRBAYES as this program incorporates the Metropolis-coupled Markov Chain Monte Carlo (MCMC) allowing for improved exploration of parameter space. The best-fit nucleotide substitution model for each partition was selected using the Akaike information criterion (AIC) implemented in JMODELTEST 0.1.1 using the three substitution scheme option (Guindon & Gascuel 2003; Posada 2008). For the mtDNA data sets, preliminary analyses were run to assess whether a codon partitioned model for the mtDNA data was a better fit to the data than using an unpartitioned model using the Bayes Factor test (Suchard *et al.* 2001) implemented in Tracer 1.5 (Drummond & Rambaut 2007). The strength of support from Bayes Factors was interpreted using guidelines from Jeffreys (1961) and Raftery

Table 2 MtDNA and nuDNA diversity statistics arranged by drainage from north to south

| mtDNA | | | | | nuDNA | | | | | | |
|---------------------|--------------------|-------------------|--------|------|-------|---------------------|--------------------|----------------|--------|------|---|
| Drainage (isolate) | No. of individuals | No. of haplotypes | π | Hd | s | Drainage (isolate) | No. of individuals | No. of alleles | π | Hd | s |
| Salt River (R) | 19 | 1 | 0.0000 | 0.00 | 0 | Salt river (R) | 17 | 3 | 0.0002 | 0.22 | 2 |
| Gila River (R) | 24 | 5 | 0.0042 | 0.72 | 12 | Gila River (R) | 11 | 2 | 0.0001 | 0.15 | 4 |
| Rio Yaqui (U) | 15 | 4 | 0.0014 | 0.54 | 6 | Rio Yaqui (U) | 13 | 10 | 0.0011 | 0.87 | 5 |
| Rio Conchos (U) | 6 | 3 | 0.0050 | 0.73 | 11 | Rio Conchos (U) | 5 | 4 | 0.0008 | 0.80 | 4 |
| Rio Fuerte (U) | 10 | 5 | 0.0324 | 0.82 | 71 | Rio Fuerte (U) | 9 | 3 | 0.0005 | 0.61 | 2 |
| Rio San Lorenzo (U) | 8 | 2 | 0.0008 | 0.29 | 3 | Rio San Lorenzo (U) | 7 | 6 | 0.0009 | 0.85 | 4 |
| Rio Quebrada(N) | 10 | 3 | 0.0004 | 0.38 | 2 | Rio Quebrada(N) | 8 | 2 | 0.0003 | 0.46 | 1 |

Each drainage is followed by an R, U or N which refer to the three geographic isolates '*rufipunctatus*', '*unilabialis*' and '*nigronuchalis*', respectively. π is nucleotide diversity; Hd is haplotypic diversity; and s is the number of polymorphic sites. Estimates do not include the individuals sequenced for Rio Nazas and Rio Casa Grandes because of limited sample sizes, also we only used individuals for which all three nuclear loci were sequenced for estimates of diversity.

(1996) with strong evidence in favour of one hypothesis over the other determined by a 2ln Bayes factor >10. The nuclear loci were each modelled separately. To assess whether the recovered clades were congruent across loci, Bayesian analyses were performed separately for the mtDNA and nuDNA loci and a concatenation of the nuclear loci. For the mtDNA, MCMC searches of tree space included four independent runs with four Markov chains each with default heating values and were run for 20 million generations. Trees were sampled every 1000th generation resulting in 20 000 trees from each run. The same scheme was followed for the nuDNA analyses except the four independent runs were estimated for 5 million generations with every 100th sample retained which resulted in 50 000 trees from each run. To concatenate the nuclear loci, we chose one of the phased nuclear gene copies at random for each individual to generate one concatenated sequence per individual. We sought evidence for convergence by visually inspecting the cumulative posterior probabilities of split frequencies using the *cumulative* and *compare* commands in the program AWTY (Nylander *et al.* 2008). Stationarity was assumed to have occurred when the cumulative posterior probabilities of all splits stabilized. Posterior probabilities (Pp) with ≥ 0.95 were considered strongly supported.

Gene trees were rooted with additional sequence data collected from *Thamnophis validus*, *Thamnophis melanogaster* and *Thamnophis scaliger* which are representatives of a clade that is most closely related to the *T. rufipunctatus* complex, as determined by a recent phylogeny of North American gartersnakes (de Queiroz *et al.* 2002).

Molecular dating and species tree estimation

To assess whether lineage diversification was consistent with Pleistocene glacial cycles or older events, we employed the species tree ancestral reconstruction option (*BEAST) in BEAST v1.6.1 (Heled & Drummond 2010), given that the common approach of concatenating sequences from multiple loci (mtDNA and nuDNA) can often lead to an incorrect but well-supported topology (Kubatko & Degnan 2007). In addition, discrepancies between gene trees and species trees have been shown to be especially problematic for closely related species (Pamilo & Nei 1988), and divergence dating based on gene trees can lead to overestimates of divergence times (Edwards & Beerli 2000; McCormack *et al.* 2010; Brandley *et al.* 2011). The distinct advantage of using *BEAST over other methods is that it allows joint inference of the species tree topology, divergence times and gene trees from the mtDNA and nuclear loci sampled from across the *T. rufipunctatus* species complex. The method uses a multispecies coalescent model,

where each gene tree is embedded inside the species tree following the coalescent back in time from the present across each branch. As such, any discrepancies between the gene trees and the species tree are attributed to incomplete lineage sorting, assuming that significant gene flow between species does not exist (an assumption we investigated with IMA2, see below). The *BEAST analysis used a yule tree prior for the species tree, employing all four loci (mtDNA and nuDNA), partition-specific substitution models used in the gene tree analyses, and every individual sequenced from the three geographic isolates, including both allelic phases of homozygous and heterozygous individuals.

Because the monophyly of the '*unilabialis*' isolate was not equally resolved between the mtDNA and the nuDNA gene tree analyses (see Results), we tested whether the multilocus species tree supported a monophyletic relationship. This was accomplished by constructing a trait table in which '*unilabialis*' individuals were separately labelled according to the northern and southern subclades that were recovered in the concatenated nuDNA analysis (except NAZ was included in the southern subclade). All other individuals were labelled according to their respective geographic isolate (i.e. '*rufipunctatus*' and '*nigronuchalis*'). Results for the *BEAST analysis were obtained by combining log and tree files from four separate runs of 40 million generations each with every 1000th sample retained using the program LOGCOMBINER, distributed as part of the BEAST package. As in other BEAST analyses, convergence statistics were monitored in AWTY.

Estimating divergence times based on fossil calibrations is more accurate if multiple calibrations spanning different parts of the tree can be incorporated (Sanders & Lee 2007; Ho & Phillips 2009). However, directly using fossil calibrations is not possible within the *T. rufipunctatus* complex because of the paucity of available data. Therefore, we expanded the mtDNA and nuDNA data sets to place calibration priors on older clades within a larger natracine phylogeny (the subfamily that includes the *T. rufipunctatus* complex). This required sequencing several additional out-groups for each nuclear locus for two individuals (when possible) per species: *Thamnophis chrysocephalus*, *T. cyrtopsis*, *T. eques*, *T. hammondii*, *T. melanogaster*, *T. pulchrilatus*, *T. sauritus*, *T. scaliger*, *T. sirtalis*, *T. validus*, *Nerodia fasciata*, *Storeria dekayi*, *Natrix maura*, *N. natrix* and *N. tessellata* (JF974262-JF974283; see Appendix I). The four temporally spaced node constraints used were based on a composite of natracine studies (Alfaro & Arnold 2001; de Queiroz *et al.* 2002; Guicking *et al.* 2006) and were previously used to estimate divergence times in a related *Thamnophis* species (in de Queiroz & Lawson 2008): (i) the division between *N. maura* and *N. na-*

trix/tesselata at 18–27 Ma (Guicking *et al.* 2006; Ma = million years ago); (ii) the division between *N. natrix* and *N. tesselata* at 13–22 Ma (Guicking *et al.* 2006); (iii) the first fossil appearance of the tribe *Thamnophini* (Holman 2000) at 18.8–19.5 Ma (Tedford *et al.* 2004); and (iv) the earliest known specimen of the monophyletic genus *Thamnophis* (Holman 2000) at 13.4–14.0 Ma (Tedford *et al.* 2004). For each age prior, we employed a lognormal distribution with an offset at the lower bound of each fossil age boundary and a standard deviation so that 95% of the distribution falls within the given age boundary. To assess potential error toward younger dates and for comparative purposes, we also performed analyses using a normal distribution and standard deviation spanning the age boundary. Results revealed very similar divergence estimates with means for all nodes differing <8.5% between the two analyses.

Coalescent estimates of gene flow

We used the isolation-with-migration model implemented in the program IMA2 (Hey & Nielsen 2004, 2007) to test whether shared haplotypes between geographic isolates were because of postisolation gene flow (i.e. introgression) or whether this pattern could be explained by retention of ancestral polymorphisms. The coalescent-based method uses MCMC simulations of gene genealogies to co-estimate various demographic parameters such as current effective and ancestral population sizes, migration rates and time of population splitting. Because the isolation-with-migration model does not assume that mutation, drift and migration are in evolutionary equilibrium, the method has a distinct advantage over others for estimating parameters for recently separated populations that may share haplotypes owing to a combination of past gene flow and ancestral polymorphism.

Recently, the method has been extended in the program IMA2 to infer demographic parameters for multiple populations (i.e. three or more; Hey 2010). We used this model to examine the population history of the three geographic isolates. Obtaining useful parameter estimates for three or more populations generally requires having data sets with numerous, highly polymorphic loci (i.e. more than the four used in this study). To help mitigate for this, we used the $-j3$ option, which limits the number of estimated migration parameters to the sampled populations only (i.e. ancestral populations have zero migration). For all analyses, we used the HKY mutation model of nucleotide substitution. To rescale results into units of time, we used the geometric mean of the mutation rates of the four loci used in this study. The mutation rate estimates for each gene were

obtained from our BEAST divergence dating analyses (multiplied by the length of the gene). We assumed a generation time of 4.5 years for *T. rufipunctatus* to obtain a measure of mutation rate on a scale of generations (Rosen & Schwalbe 1988; P. C. Rosen, personal communication). We used a series of initial runs to determine the most appropriate search parameters that maximized mixing, using broad maximum priors for population parameters (i.e. m , θ , and t). Following these initial runs, a long analysis was run with final priors (see below) to monitor for convergence and to determine the appropriate burn-in (750 000 steps). The state of the Markov chain was saved from this analysis and used to seed four separate runs (30 000 genealogies saved from each). Final priors used for each analysis included 40 chains with a geometric heating scheme ($h_a = 0.975$, $h_b = 0.75$), maximum migration prior value ($m = 3$), maximum population size parameter ($q = 8$) and a maximum time of population splitting prior ($t = 4$). These runs were combined to generate 120 000 gene trees for use with likelihood ratio tests implemented in the nested models option in 'Load Trees Mode'. Likelihood ratio test were used to help determine whether models that allowed fluctuations in migration rate were a better fit to the data than a simpler model with zero gene flow.

Results

Demographic and genetic diversity analyses

HKA tests for each locus did not reveal any deviation from neutral expectations (P -values were >0.2 across all loci). According to the DSS test, only the R35 locus showed evidence of recombination between positions 375 and 422. Therefore, this portion of sequence was trimmed from the data before being used in subsequent analyses. Genetic diversity estimates (π) within drainages ranged from 0.0000 to 0.0324 and 0.0000 to 0.0017 for mtDNA and nuDNA, respectively (Table 2). In general, drainages with the lowest diversity estimates were found along the northern and southern range limits, while drainages along the interior had higher nucleotide diversity. Regressions of nucleotide diversity (π) estimates from both mtDNA and nuDNA vs. latitude were not significant (mtDNA: $R^2 = 0.031$, slope = 0.000, $P > 0.675$; nuDNA: $R^2 = 0.291$, slope = -0.005 , $P > 0.167$) suggesting that diversity does not vary with latitude. Nonetheless, levels of nucleotide diversity are not equal across geographic isolates. Nucleotide diversity assessed across all loci is more than 10-fold higher along the Sierra Madre ('*unilabialis*' isolate) than it is in the smaller geographic isolates (Table 3). These results suggest that the area of each isolate is a better determi-

Table 3 MtDNA and nuDNA diversity statistics and average divergence for each isolate, and pairwise Φ_{CT} values

| Locus | Length (bp) | Nucleotide diversity, π | | | | Average divergence | | | Φ_{CT} | |
|----------------|-------------|-----------------------------|--------|--------|--------|--------------------|---------|---------|-------------|--------|
| | | R | U | N | Total | R-U (%) | U-N (%) | R-N (%) | R-U | U-N |
| NADH1 | 1090 | 0.0034 | 0.0275 | 0.0004 | 0.0275 | 2.24 | 4.68 | 5.77 | 0.2261 | 0.3455 |
| Combined nuDNA | 1705 | 0.0002 | 0.0021 | 0.0003 | 0.0026 | 0.40 | 0.39 | 0.39 | 0.6067 | 0.5216 |

Group abbreviations R, U, N correspond to the '*rufipunctatus*', '*unilabialis*' and '*nigronuchalis*' identified in Fig. 4. Average length of each locus in base pairs; within clade diversity, π ; average sequence divergence between groups; pairwise Φ_{CT} among groups.

nant of genetic diversity than latitude. Patterns of mtDNA variability revealed higher average divergence between comparisons with the '*nigronuchalis*' isolate than between the '*rufipunctatus*' and '*unilabialis*' comparison (Table 3), whereas the concatenated nuDNA estimates revealed similar divergences across all comparisons, but all loci showed significant Φ_{CT} values among isolates. Results from the AMOVA comparisons revealed that most of the nuDNA genetic variation can be explained among the geographic isolates or three species hypothesis (65.88%) rather than the two species hypothesis (44.94%), while mtDNA genetic variation revealed similar proportions of variation across both comparison (58.86% and 54.07%, respectively; Table 4).

Phylogenetic gene tree analyses

A total of 1098 bp of mtDNA (148 bp 16S rRNA and 964 bp ND1) were collected. Eight of the 16S rRNA characters were excluded from the analyses because of uncertain alignment. Among the remaining 1090 aligned nucleotide positions, 96 were variable among *Thamnophis rufipunctatus* complex sequences and 92 were parsimony informative. Six unique haplotypes were recovered from the 43 '*rufipunctatus*' specimens, 14 unique haplotypes from the 40 '*unilabialis*' specimens and three unique haplotypes from the 10 '*nigronuchalis*'

specimens sampled. Bayes factors strongly suggest that the partitioned model is a better fit to the mtDNA data (2Ln Bayes Factor = 321). Three different models of nucleotide substitution were identified using AIC: GTR for the 16S region, HKY + I for ND1-first and ND1-second codon positions, and GTR + I for the ND1-third codon position. The mtDNA Bayesian analysis provided strong support for three major clades (posterior probabilities of 0.99–1.0) along a north-to-south axis, although haplotypes of each clade did not always come from the same geographic isolate (Fig. 2). In particular, the northern clade formed a polytomy composed of both '*rufipunctatus*' and '*unilabialis*' specimens that are distributed along the Mogollon Rim within the Gila and Salt drainages and within drainages along the northern portion of the Sierra Madre Occidental down to the headwaters of the Rio Conchos (BAL). The Central clade is composed of the remaining '*unilabialis*' specimens that are distributed within drainages along the Sierra Madre Occidental starting at the southern headwaters of the Rio Fuerte (VER) southward to the Rio San Lorenzo (VUE, NIC, QUE) and Rio Nazas (NAZ). A Southern clade, composed entirely of '*nigronuchalis*', is geographically isolated from the previous two clades and occurs within the Rio del Presidio drainage. Therefore, according to the topology of mtDNA Bayesian consensus tree, the '*unilabialis*' isolate is polyphyletic,

Table 4 MtDNA and nuDNA hierarchical analysis of molecular variance (AMOVA) results for the *Thamnophis rufipunctatus* complex

| Groups | No. of groups | Source of variation | % of variation (mtDNA) | P-value | % of variation (nuDNA) | P-value |
|--|---------------|-----------------------------------|------------------------|---------|------------------------|---------|
| Geographic isolates (three species hypothesis) | 3 | Among isolates | 43.89 | 0.0061 | 65.88 | <0.000 |
| | | Among populations within isolates | 54.07 | <0.000 | 21.34 | <0.000 |
| | | Within populations | 2.04 | <0.000 | 12.79 | <0.000 |
| Species (two species hypothesis) | 2 | Among species | 58.86 | 0.0431 | 44.94 | 0.024 |
| | | Among populations within species | 39.85 | <0.000 | 44.69 | <0.000 |
| | | Within populations | 1.29 | <0.000 | 10.38 | <0.000 |

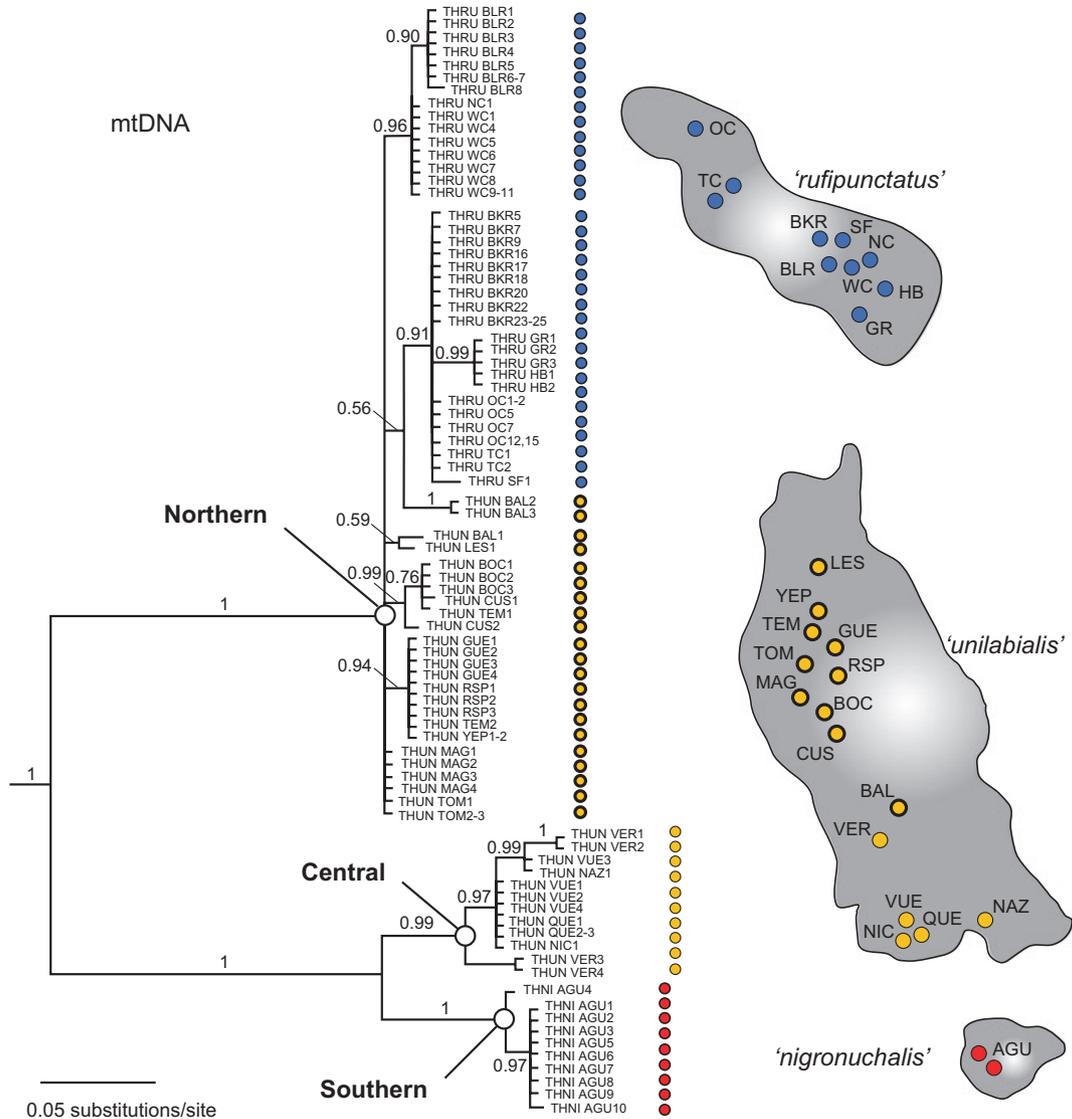


Fig. 2 Phylogenetic estimate for the *Thamnophis rufipunctatus* species complex based on partitioned Bayesian analysis of mitochondrial DNA sequence data. Site localities for the three geographic isolates '*rufipunctatus*', '*unilabialis*' and '*nigronuchalis*' are coloured blue, orange and red, respectively. Orange site localities outlined with a thick black line highlight the instances of discordance between mitochondrial and nuclear data. The numbers on branches indicate Bayesian posterior probabilities (P_p). Outgroups have been pruned from the phylogram to improve clarity.

with southern '*unilabialis*' being more closely related to the '*nigronuchalis*' clade ($P_p = 1.0$), while northern '*unilabialis*' populations are interspersed within a clade containing '*rufipunctatus*' (Fig. 2). Among these mtDNA clades, the highest level of sequence divergence is between the northern and central clades at 5.7% and lowest between the central and southern clades at 2.3%.

For the nuDNA gene trees, we collected a total of 1705 bp from the three nuclear loci across the *T. rufipunctatus* species complex. The total number of heterozygous individuals per locus was 13 of 84 for TATA, 3 of 82 for VIM5 and 24 of 82 for R35. The best-fit substi-

tution models identified by AIC were HKY for the VIM5 locus and HKY + I for both TATA and R35 loci. Bayesian analyses of individual nuDNA gene trees exhibited differing degrees of congruence with each other (Fig. S1, Supporting information). In general, the R35 and VIM5 gene trees revealed clusters of alleles that were geographically congruent to the three previously proposed taxonomic units, while the TATA gene tree revealed greater admixture among the '*unilabialis*' and '*nigronuchalis*' isolates. Across all three loci, the '*rufipunctatus*' isolate exhibited allelic variation primarily exclusive from all other populations. Unlike the

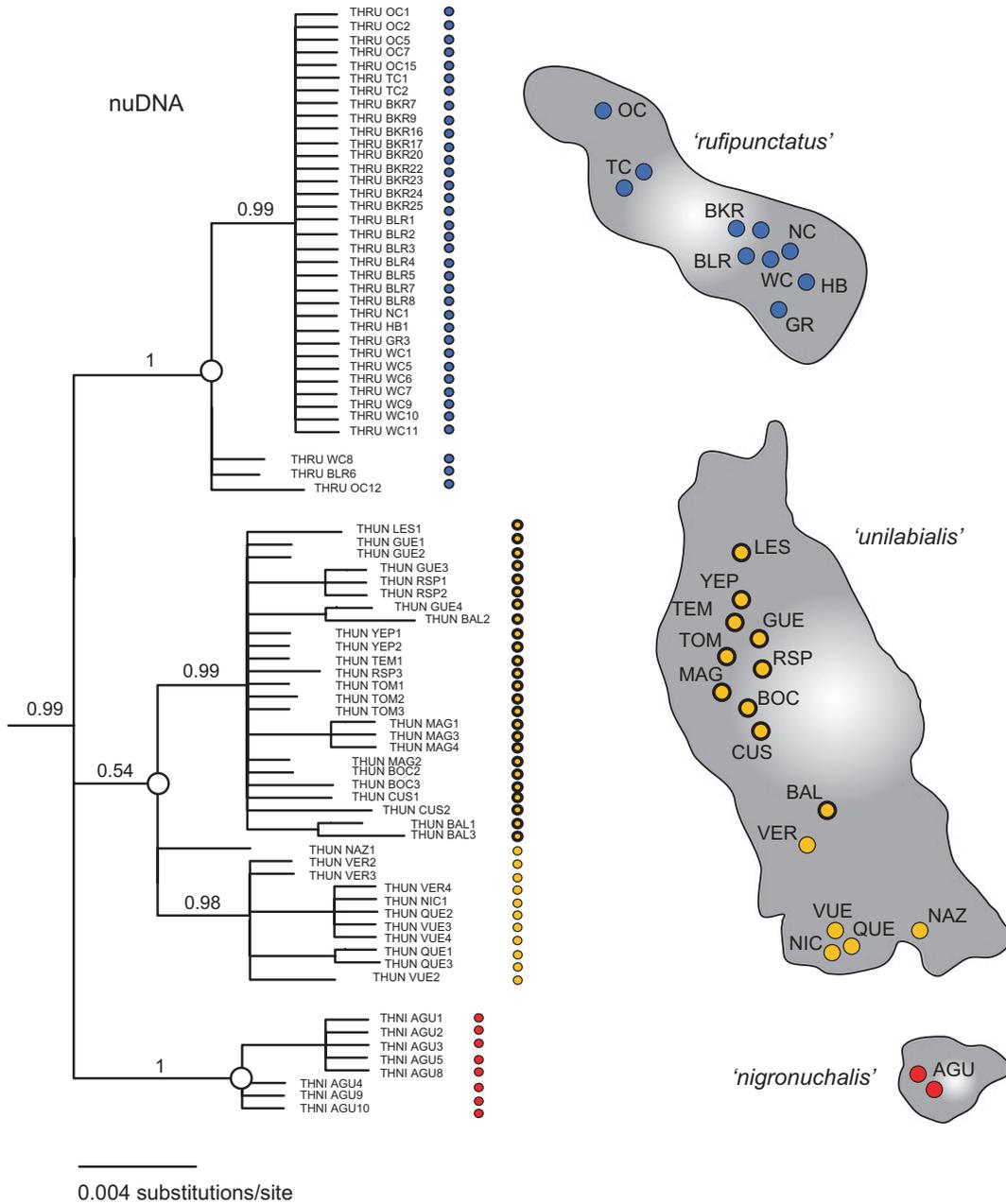


Fig. 3 Phylogenetic estimate of the *Thamnophis rufipunctatus* species complex based on Bayesian analysis of concatenated nuclear sequences from TATA, R35 and VIM5. Site localities for the three isolates '*rufipunctatus*', '*unilabialis*' and '*nigronuchalis*' are coloured blue, orange and red, respectively. Orange site localities outlined with a thick black line highlight the instances of discordance between mitochondrial and nuclear data. The numbers on branches indicate Bayesian posterior probabilities (P_p). Outgroups have been pruned from the phylogram to improve clarity.

mtDNA analysis, phylogenetic relationships inferred from the concatenated nuDNA Bayesian analysis supported three major clades that are consistent with present-day distributions of the geographic isolates ('*rufipunctatus*', '*unilabialis*' and '*nigronuchalis*'; Fig. 3). However, the monophyly of the '*unilabialis*' clade is weakly supported ($P_p = 0.54$) and is composed of two subclades (although NAZ falls outside these subclades)

that are geographically subdivided between populations BAL and VER, the same location where northern and central mtDNA clades are partitioned.

Divergence-times and species tree estimate

Bayes factor comparisons decisively favoured an uncorrelated lognormal relaxed clock rate over the strict clock

(2Ln Bayes Factors >100). Each of the four *BEAST analyses that were run with calibration age constraints achieved stationarity between 2 and 3 million generations. We discarded the first 4 million generations (20%) from each analysis as a conservative burnin. Thus, all subsequent clade posterior probabilities and divergences-time estimates were pooled from the remaining 128 000 postburnin trees. The multi-locus species tree strongly supports the placement of the southern most isolate ('*nigronuchalis*') as the sister lineage to all other isolates (Fig. 4). This is followed by a paraphyletic relationship between northern and southern '*unilabialis*' lineages with respect to '*rufipunctatus*'. Although the multi-locus species tree does not resolve the '*unilabialis*' isolate as monophyletic, paraphyly is only weakly supported ($P_p = 0.54$). In general, divergence-time estimates support an initial lower Pleistocene diversification (1.51 Ma; Fig. 4) for the *Thamnophis rufipunctatus* species complex.

IMA: Gene flow and divergence estimation

Under the multi-population isolation-with-migration model, asymmetric migration rates were recovered from

the '*unilabialis*' isolate into both the '*rufipunctatus*' and the '*nigronuchalis*' isolates (Table 5), while all other migration estimates were either non-significant and/or zero. Likelihood ratio tests comparing alternative nested coalescent migration models revealed that models assuming equal and zero migration between geographic isolates could not be rejected (Table 6). These results indicate that a no migration model is a better fit to the data than the full migration model. Divergence time estimates among '*nigronuchalis*' and all other isolates converged at 530 ka (ka = thousand years ago, 95% highest posterior density (HPD) ranging from 186 to 995 ka), while slightly younger divergence time estimates were recovered among '*unilabialis*' and '*rufipunctatus*' converging at 141 ka (95% HPDs ranging from 64 to 298 ka). However, both estimates fall entirely within the Pleistocene. Demographic estimates of effective population sizes scaled by the generation time were largest in the '*unilabialis*' isolate and much smaller in both '*rufipunctatus*' and '*nigronuchalis*' isolates (Table 5).

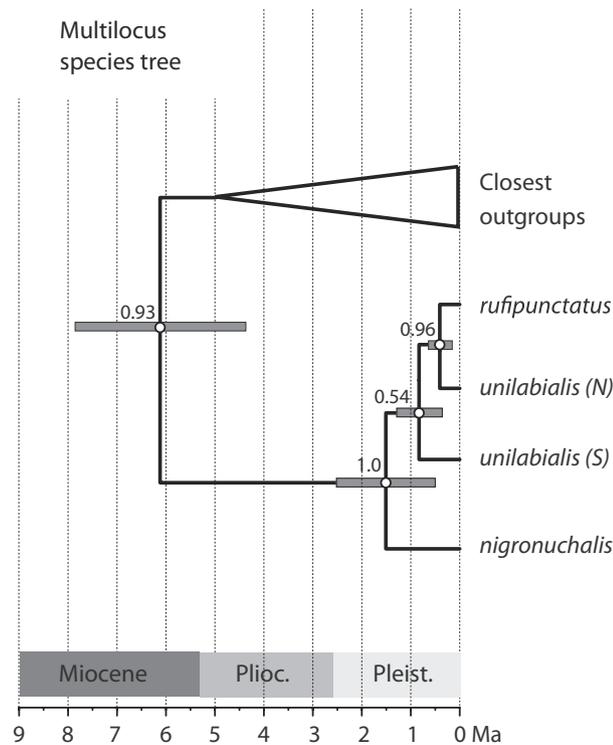


Fig. 4 Chronogram and divergence time estimates from multi-locus species tree analysis for the *Thamnophis rufipunctatus* species complex. The chronogram was trimmed to exclude most outgroup taxa for clarity purposes. The number at each node represents the posterior probability support value for the node, and grey bars indicate the 95% HPDs of the node age.

Table 5 Demographic parameter estimates from the multi-population isolation with migration model of the *Thamnophis rufipunctatus* complex

| Parameter | Point estimate | 95% HPD (low/high) |
|-------------------------------------|----------------|-----------------------|
| Population migration rate $2N_m$ | | |
| $2N_{m0} > 1$ (R ← U) | 0.076* | 0.005/0.259 |
| $2N_{m1} > 0$ (U ← R) | 0.106 | 0.000/0.450 |
| $2N_{m1} > 2$ (U ← N) | 0.001 | 0.004/0.252 |
| $2N_{m2} > 1$ (N ← U) | 0.048* | 0.012/0.254 |
| $2N_{m0} > 2$ (R ← N) | 0.000 | 0.000/0.095 |
| $2N_{m2} > 0$ (N ← R) | 0.001 | 0.000/0.107 |
| Effective population size (N_e) | | |
| N_e (R) | 12 923 | 5979/33 755 |
| N_e (U) | 91 620 | 51 886/150 643 |
| N_e (N) | 5594 | 1350/20 639 |
| Divergence estimate (t) | | |
| t_0 (R vs. U) | 140 613 years | 64 230/296 849 years |
| t_1 (R, U vs. N) | 529 468 years | 185 748/994 705 years |

Joint point demographic estimates for population pairwise comparisons: the population migration rate ($2N_m X > Y$) is interpreted forward in time and equals the rate at which genes of pop X are replaced by pop Y ($X \leftarrow Y$), effective population size (N_e), divergence estimates (t) in years, and 95% highest posterior density intervals. Point estimates for each parameter are taken from the bins with highest estimates. R, U and N refer to *rufipunctatus*, *unilabialis* and *nigronuchalis* isolates, respectively. Asterisks identify statistically significant ($P < 0.01$) migration using the test of Nielsen & Wakeley 2001.

Table 6 Log-likelihood ratio tests of nested coalescent migration models from the multi-population isolation with migration analysis

| Model description | Model | Degrees of freedom | -2 Δ | P |
|---------------------------|---|--------------------|-------------|-------|
| Full migration model | $m_0 > 1$ $m_1 > 0$ $m_0 > 2$ $m_2 > 0$ $m_1 > 2$ $m_2 > 1$ | — | — | — |
| Equal migration rates | $m_0 > 1 = m_1 > 0$ $m_0 > 2$ $m_2 > 0$ $m_1 > 2$ $m_2 > 1$ | 5 | 9.49 | 0.09 |
| All migration set to zero | $m_0 > 1$ $m_1 > 0$ $m_0 > 2$ $m_2 > 0$ $m_1 > 2$ $m_2 > 1 = 0$ | 6 | 8.49 | 0.204 |

Migration parameter descriptions are as follows: m_0 equals the migration rate for '*rufipunctatus*', m_1 equals the migration rate for '*unilabialis*', m_2 equals the migration rate for '*nigronuchalis*'. The -2Δ value is the calculated chi-square test statistic between the full and the nested migration models. Significance was assessed at $P < 0.05$.

Discussion

Phylogeography-single southern vs. multiple refugia

In this study, we first examined whether patterns of interpopulation genetic divergence support one of the two models of range fluctuation: (i) an archipelago model of range fluctuation consistent with the present-day geographic isolates observed across the range of the *Thamnophis rufipunctatus* species complex (multiple refugia) or (ii) the latitudinal shift model represented by a single southern refugium that experienced northward expansion. Specifically, predictions of the archipelago model suggest lineage diversity is driven by expansion and contraction of suitable habitats along elevation gradients resulting in multiple centres of genetic diversity across the species distribution. Therefore, the distribution of divergent lineages supported by the genetic data should reflect the distribution of montane isolates. Also genetic diversity and coalescent times across these isolates should not be strongly correlated with the latitude. In contrast, the latitudinal model predicts decreasing genetic variation and coalescent times with increasing latitude across the species distribution consistent with dispersal out of southern refugia.

Our data provide evidence for multiple lineages that correspond to present-day geographic isolates (and taxonomic units), implying that the *T. rufipunctatus* species complex was sundered in multiple refugia during the Pleistocene. Our divergence-time estimates show that the impact of the Pleistocene climate cycles was readily apparent in each geographic isolate (rather than older Neogene events), although it must be noted that the recovered coalescent times for divergence of the '*nigronuchalis*' isolate (the most southern isolate) were deeper (>500 ka; Table 5, Fig. 4). Diversity estimates across the major drainages did not decrease with increasing latitude as expected under the latitudinal shift model. Instead, higher diversity and largest effective population sizes (N_e) were recovered centrally along the Sierra Madre Occidental, while substantially smaller effective population sizes and genetic diversity estimates were

recovered in the northern and southern peripheral isolates.

Taken together, our results are most consistent with the archipelago model of range fluctuation involving Pleistocene diversification of multiple refugia. Similar to the mechanisms hypothesized for more northern sky island species of North America (e.g. Hewitt 1996; Knowles 2000; DeChaine & Martin 2006), these refugia likely evolved within and across the separate mountain systems through reoccurring episodes of up-slope contraction during interglacials and down-slope expansion during cooler glacial periods. Even though contact between isolates might have occurred along mountain corridors during cooler glacial periods, genetic exchange must have been limited to the lower-elevation populations that existed between major refugia given the topographic heterogeneity of the region. As has been suggested for other high-elevation taxa (McCormack *et al.* 2008b; Galbreath *et al.* 2009), these ephemeral low-elevation populations were probably the first to be extirpated as populations tracked well-watered surface environments to higher elevations during warmer interglacials and evidence of genetic exchange would have been significantly reduced and/or erased via genetic drift and lineage sorting. In addition, historical disjunctions recovered within the Sierra Madre Occidental (e.g. multiple '*unilabialis*' clades) suggest geographic barriers, such as the deep canyons and extensive drainage systems that empty into coastal and interior basins of the Mexican Plateau, may have played a role in disrupting glacial corridors as well. Based on the multilocus divergence estimates, the most recent divergence among geographic isolates of the *T. rufipunctatus* complex involves '*rufipunctatus*' and '*unilabialis*'. This isolation event was preceded by the separation of the ancestor of these two isolates and '*nigronuchalis*'. We hypothesize that ancestral populations were likely centred along the main stem of the Sierra Madre Occidental as our data reveal highest diversity, as well as largest N_e and higher signals of historical gene flow from this isolate to those in the north and south. Subsequent population expansions and contractions associ-

ated with episodes of glacial cooling and interglacial warming likely provided different opportunities for peripatric lineage diversification (episodes of budding off) via separate southward and northward (in that order) founder events from the main stem of the Sierra Madre Occidental. In addition to geographic separation, it is possible that selection may have enhanced differentiation among the lineages (Endler 1986; Nielsen 2005), as suggested by maintenance of divergent phenotypes found across the three isolates (Thompson 1957; Rossman 1995).

MtDNA vs. nuDNA genealogies-incomplete lineage sorting or introgression?

Phylogenetic analyses of gene trees identified three divergent clades; however, the geographic boundaries of these clades differed between mtDNA and nuDNA analyses. Specifically, the nuDNA analyses revealed clades ('*rufipunctatus*', '*unilabialis*' and '*nigronuchalis*') that coincide with the three geographic isolates and previously named taxa, although support for a monophyletic '*unilabialis*' clade was low. In contrast, mtDNA showed northern populations of '*unilabialis*' interspersed within a clade also containing '*rufipunctatus*'. This mtDNA pattern was recovered despite the fact that these isolates are currently separated by *c.* 350 km of unsuitable lowland habitat. Such cases of discordance among mtDNA and nuDNA gene trees are typically attributed to incomplete lineage sorting of ancestral haplotypes, recent mitochondrial introgression or differences between male and female dispersal rates (Ballard *et al.* 2002; Funk & Omland 2003; Ballard & Whitlock 2004).

Greater female dispersal capabilities are unlikely, as preliminary evidence for US populations of *T. rufipunctatus* suggests a lack of differential movement and home range size between males and females (Nowak 2006; Jennings & Christman 2009). While it is often difficult to differentiate between incomplete lineage sorting and introgression (Funk & Omland 2003), several lines of evidence suggest the observed incongruence is largely driven by incomplete sorting of recently separated lineages. First, if incomplete lineage sorting is driving the observed mtDNA pattern, similar nonexclusive patterns should also be expected in the nuclear genes given the potentially larger effective population size for nuDNA and hence a longer sorting time (Moore 1995). The results from individual nuDNA gene tree analyses (Fig. S1, Supporting information) are consistent with this expectation. Each nuDNA gene tree exhibited differing degrees of congruence and resolution with respect to each other and the three geographic isolates, which emphasizes the stochastic variance of genetic

processes. However, when these loci were combined in the concatenated nuDNA analyses (Fig. 3) three geographically cohesive clades, consistent with the three geographic isolates, were recovered implying that phylogenetic signal accumulated through the addition of multiple loci. These results echo several recent studies that have shown how better genealogical resolution can be gained through the addition of multiple nuclear markers, rather than relying solely on a single locus (reviewed in Edwards & Bensch 2009). A second line of evidence in support of incomplete lineage sorting comes from our divergence time estimates and species tree history. All things being equal, the time after isolation at which a pair of populations (or species) is observed will be an important factor influencing their genealogical status (i.e. polyphyletic to monophyletic lineages). Our data provide evidence for more recent Pleistocene differentiation of these northern isolates (<1 Ma) rather than older Neogene events indicating that time as divergence has not been sufficiently long for the corresponding '*rufipunctatus*' and '*unilabialis*' lineages to fully sort. (Fig. 4, Table 5). Finally, by analyzing the four independent loci with an isolation-with-migration coalescent model, we were able to specifically test whether incomplete lineage sorting because of recent divergence could explain topological discrepancies observed between the mtDNA and the nuDNA gene lineages or whether past non-zero levels of gene flow (i.e. introgression) are necessary to explain the observed phylogenetic patterns. Based on the joint estimates of demographic parameters (i.e. migration and divergence time), a zero gene flow model and more recent divergence times (with respect to '*nigronuchalis*') were supported between the '*rufipunctatus*' and '*unilabialis*' isolates. Therefore, our approach provided a quantitative framework to better understand causes of gene tree discordance between this closely related species complex. While the large geographical distance separating related haplotypes of '*rufipunctatus*' from those of '*unilabialis*' would seem to preclude any gene flow between these geographic isolates, our genetic results suggest some allele mixing among these isolates is a legacy of historical migration that occurred prior to habitat fragmentation and isolation.

Taxonomy and conservation

Accurate delimitation of species diversity is increasingly important, as species distributions are becoming progressively more reduced and threatened. Understanding the genetic diversity and differentiation of populations within species is at the heart of phylogeographic studies. However, unravelling and accurately describing this diversity at the early stages of speciation is often chal-

lenging, because a number of molecular phenomena may obfuscate organismal lineages (e.g. ancestral polymorphism, introgression, selection). Ideally, evaluating diversity across multiple operational criteria (genetic, ecological and morphological) provides the most robust evidence to delimit species and increases the ability to detect recently separated lineages. For the *T. rufipunctatus* species complex, limited data are available on species ecology, habitat usage and morphological differentiation among the currently defined taxonomic classifications. As such, our genetic results help to identify lineage separation and divergence across this species group and provide an initial framework for taxonomic and conservation management decisions.

Taken together, our analyses identify lineages concordant with previously recognized taxonomic divisions and geographic isolates (i.e. *rufipunctatus*, *unilabialis* and *nigronuchalis*) indicating that three separate species may exist within the *T. rufipunctatus* complex. Rossman (1995) supported recognition of *Thamnophis nigronuchalis* as a distinct species on the basis of his morphological studies and molecular results published elsewhere (de Queiroz & Lawson 1994). Morphological differences and the allopatric distribution of *T. nigronuchalis* suggested absence of gene flow, and our data support this conclusion. We also recovered evidence of diversification and zero gene flow between 'unilabialis' and 'rufipunctatus' isolates. However, whereas both the mtDNA and nuDNA genealogies identify a discrete 'nigronuchalis' lineage, topological patterns of divergence between 'rufipunctatus' and 'unilabialis' among gene trees were not immediately clear (Figs 2 and 3). Reinterpreting these relationships in the light of the multi-locus species analyses provided evidence of a paraphyletic history for 'unilabialis' with respect to 'rufipunctatus' (Fig. 4). Considering the relatively small estimate of N_e for 'rufipunctatus' coupled with the general geographic context of 'rufipunctatus' in relation to 'unilabialis', a peripatric mode of speciation for these morphologically diagnosable sister taxa seems consistent. Additionally, if we treat northern 'unilabialis' as a monophyletic genetic subgroup of 'unilabialis' and consider the inferred absence of gene flow and phenotypically divergent characters (muzzle length, eye diameter, postrostral and ventral scale counts; Rossman 1995) between 'unilabialis' and 'rufipunctatus', the evidence would suggest that these geographic isolates are on independent evolutionary trajectories and warrant recognition as full species: the Mogollon Narrowheaded Gartersnake, *T. rufipunctatus* (Cope) and the Madrean Narrowheaded Gartersnake, *T. unilabialis* (Tanner). That said, further hypothesis testing might be warranted with additional morphological and/or ecological data sets to assess the adaptive significance between these distinct lineages

(Crandall *et al.* 2000). Future morphometric analyses that quantify differences of ecologically relevant characters (e.g. head and snout morphology) attributed to underwater foraging adaptations may help to clarify a greater degree of concordance between morphology and genetic divergence identified in this study.

Our study has further conservation and management implications for *T. rufipunctatus* (*sensu novum*) that can help mitigate loss of genetic diversity in an already threatened genetic landscape. Recent surveys along the two drainage systems (Gila and Salt Rivers) occupied by these snakes indicate remaining populations are geographically isolated, declining and/or experiencing local extirpations (Rosen & Schwalbe 1988; Nowak & Santana-Bendix 2002; Holycross *et al.* 2006). One of the central management concerns has been whether these disjunct Mogollon Rim populations correspond to a distinct genetic group. First, our analyses demonstrate that this isolate represents a distinct evolutionary lineage and from a management perspective would merit conservation recognition (even if not recognized as a species) under several definitions (Dizon *et al.* 1992; Moritz 1994; Crandall *et al.* 2000; Fraser & Bernatchez 2001). Second, by sampling throughout the range of *T. rufipunctatus* (*sensu novo*), we were able to define population genetic structure and diversity across all known extant populations in the US and showed that phylogeographic patterns are largely congruent with contemporary drainage patterns. Based on our diversity analyses, populations along the Salt (Salt, Black, Tonto and Verde rivers) and Gila (upper Gila, San Francisco and Blue rivers) drainages harbour similarly low levels of diversity (Table 2), although the Salt River populations harbour almost twice the nuDNA diversity as the Gila populations. Given the small size of populations, their apparent isolation from each other, and the rapidity with which populations are disappearing, *T. rufipunctatus* (*sensu novo*) appears to be highly vulnerable to extinction. The urgency of this crisis and the vulnerability of populations are both exemplified by the rapid (<10 years) and mysterious disappearance of one of the most robust U.S. populations, along the San Francisco River, New Mexico (Hibbitts *et al.* 2009). Although the primary causes of range-wide declines are attributed to the introduction of non-native species and habitat alteration, additional pressures may be due to intrinsic negative genetic effects. Small effective population sizes and/or low gene flow among populations can act to reduce population viability via inbreeding depression (Saccheri *et al.* 1998; Frankham 2005) and/or reduced fitness because of increased genetic load (Hedrick & Fredrickson 2010). Empirical research and theory both show a positive link between genetic variation and fitness traits (O'Brien & Evermann 1988; Quattro & Vrijen-

hoek 1989; Hedrick & Kalinowski 2000). In addition to habitat restoration efforts for populations experiencing declines and extirpations, management and regulatory agencies should focus on maintaining maximum genetic diversity within and between drainages and consider the possibility of implementing genetic rescue efforts for populations in these drainages systems (Hedrick & Fredrickson 2010). Given the lack of diversity noted in the nuclear genes assessed here, additional studies using more variable nuclear markers (e.g. microsatellites) may be more useful to inform genetic management decisions within these U.S. drainages.

Summary and conclusions

Our results provide evidence that climatic changes during the Pleistocene had profound effects on the *T. rufipunctatus* complex lineage diversification and demography. Our nuDNA analyses recovered evidence for three divergent lineages concordant with the present-day geographic isolates, despite mtDNA evidence of ephemeral connectivity during the most recent glacial expansions. We propose that the combined effects of glacial-interglacial climate cycles along with genetic drift and selection in the smaller peripheral isolates may have enhanced isolation of these lineages throughout the warming periods of the late Pleistocene (<0.7 Ma). Our study is one of few detailed phylogeography studies of a terrestrial vertebrate inhabiting the Sierra Madre Occidental and northern sky islands of the southwestern USA (although see Bryson *et al.* 2010, 2011; McCormack *et al.* 2010). The spatial patterns of molecular variation along the Sierra Madre Occidental suggest a complex late Pleistocene history involving long periods of isolation (particularly in the south) and a demographic history of stability (particularly in the north) with more recent diversification of the more dynamic range limit in the USA. Given the region's topographically complex mountain system, climatic shifts induced by the Pleistocene may have produced differing genetic effects for other co-distributed Mexican Highland species. Yet similar patterns of isolation of southern regions (e.g. southern Durango) from more northern populations along the Sierra Madre Occidental have been recovered in Montane rattlesnakes (*Crotalus triseriatus* group; Bryson *et al.* 2011), Mexican Pine Beetles (Anducho-Reyes *et al.* 2008; *Dendroctonus mexicanus*) and Mexican Douglas-fir (*Pseudotsuga menziesii*; Gugger *et al.* 2011), suggesting that many of the deep east-west drainages that empty into the coastal plains and interior basins of the Mexican Plateau (i.e. Rio Fuerte-Conchos, Rio del Presidio, Rio Culiacán, Rio Mezquital basins) may function as effective barriers to gene flow. In addition, a recent phylogeographical study by Gugger *et al.*

(2011) on Mexican Douglas-fir (*P. menziesii*) recovered similar spatial patterns of molecular variation to our study along the Sierra Madre Occidental and strong evidence for late Pleistocene diversification of Mexican populations relative to northern U.S. populations. Given that Mexican Douglas-fir is a major component of the Madrean pine-oak woodlands of the Sierra Madre Occidental, this study provides some evidence that specialized community associations may have responded equally to the same environmental changes associated with Quaternary climate cycles. While our study and those mentioned earlier can be used to build a historical framework for understanding current geographic patterns of diversification along the Mexican highlands and sky islands of the southwestern USA, a more complete understanding of the impact of Quaternary climate cycles on divergence and local adaptation will probably come from comparative phylogeographical studies focused on ecologically associated species.

Acknowledgements

This project was funded by Arizona Game and Fish Commission Heritage Award I-08011 and the U.S. Geological Survey, Western Ecological Research Center. We thank Bill Burger, Eric Centenero Alcala, Anny Peralta Garcia, Toby Hibbitts, Ioana Hociota, Tom Jones, Mayra Moreno, Luis Oliver, Jim Rorabough, Jeff Servoss and Jorge Valdez Villavicencio for assistance in the field. For tissue loans, we thank Valerie Boyarski, Thomas Brennan, Bruce & Michelle Christman, Brian Jennings, Erika Nowak, Phil Rosen, Charlie Painter and the following institutions: Arizona State University, California Academy of Sciences (Jens Vindum), Museum of Vertebrate Zoology (Jimmy McGuire & Carol Spencer); Louisiana State University (Chris Austin & Donna Dittman), Texas Cooperative Wildlife Museum (Toby Hibbitts) and University of Arizona. Specimens and tissues collected and exported from Mexico were authorized by SEMARNAT permits (#00914/07, #01958/07, & 22736), TUFESA (#270507) and PROFEPA (#EF26-1264-07-NUG). The use of trade, product or firm names in this publication does not imply endorsement by the U.S. Government.

References

- Alfaro ME (2002) Forward attack modes of aquatic feeding garter snakes. *Functional Ecology*, **16**, 204–215.
- Alfaro ME, Arnold SJ (2001) Molecular systematics and evolution of *Regina* and the thamnophiine snakes. *Molecular Phylogenetics and Evolution*, **21**, 408–423.
- Anducho-Reyes MA, Cognato AI, Hayes JL *et al.* (2008) Phylogeography of the bark beetle *Dendroctonus mexicanus* Hopkins (Coleoptera: Curculionidae: Scolytinae). *Molecular Phylogenetics and Evolution*, **49**, 930–940.
- Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.
- Ballard JW, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

- Ballard JW, Chernoff B, James AC (2002) Divergence of mitochondrial DNA is not corroborated by nuclear DNA, morphology, or behavior in *Drosophila simulans*. *Evolution*, **56**, 527–545.
- Brandley MC, Wang Y, Guo X *et al.* (2011) Accommodating heterogeneous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon (Eumeces)* lizards. *Systematic Biology*, **60**, 3–15.
- Bryson Jr RW, Nieto-Montes de Oca A, Jaeger JR, Riddle BR (2010) Elucidation of cryptic diversity in a widespread nearctic treefrog reveals episodes of mitochondrial gene capture as frogs diversified across a dynamic landscape. *Evolution*, **64**, 2315–2330.
- Bryson Jr RW, Murphy RW, Lathrop A *et al.* (2011) Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattlesnakes. *Journal of Biogeography*, **38**, 697–710.
- Crandall KA, Beninda-Emonds ORP, Mace GM *et al.* (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, **15**, 290–295.
- DeChaine EG, Martin AP (2004) Historic cycles of fragmentation and expansion in *Parnassius smintheus* (Papilionidae) inferred using mitochondrial DNA. *Evolution*, **58**, 113–127.
- DeChaine EG, Martin AP (2006) Using coalescent simulations to test the impact of Quaternary climate cycles on divergence in an alpine plant-insect association. *Evolution*, **60**, 1004–1013.
- Dizon AE, Lockyer C, Perrin WF *et al.* (1992) Rethinking the stock concept: a phylogeographic approach. *Conservation Biology*, **6**, 24–36.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.
- Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839–1854.
- Edwards SV, Bensch S (2009) Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2009. *Molecular Ecology*, **18**, 2930–2933.
- Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.
- Flehart ED (1967) Comparative ecology of *Thamnophis elegans*, *T. cyrtopsis*, and *T. rufipunctatus* in New Mexico. *Southwestern Naturalist*, **12**, 207–230.
- Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.
- Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for identifying conservation units. *Molecular Ecology*, **10**, 2741–2752.
- Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics*, **34**, 397–423.
- Galbreath KE, Hafner DJ, Zamudio KR (2009) When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American Pika, *Ochotona princeps*). *Evolution*, **63**, 2848–2863.
- Gugger PF, Gonzalez-Rodriguez A, Rodriguez-Correa H *et al.* (2011) Southward Pleistocene migration of Douglas-fir into Mexico: phylogeography, ecological niche modeling, and conservation of 'rear-edge' populations. *New Phytologist*, **189**, 1185–1199.
- Guicking D, Lawson R, Joger U, Wink M (2006) Evolution and phylogeny of the genus *Natrix* (Serpentes: Colubridae). *Biological Journal of the Linnean Society*, **87**, 127–143.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.
- Hedrick PW, Fredrickson R (2010) Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservation Genetics*, **11**, 615–626.
- Hedrick PW, Kalinowski ST (2000) Inbreeding depression in conservation biology. *Annual Review in Ecology and Systematics*, **31**, 139–162.
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.
- Hewitt G (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Hewitt G (2004) Genetic consequences of climate oscillations in the Quaternary. *Philosophical Transactions of the Royal Society London B*, **359**, 189–195.
- Hey J (2010) The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. *Molecular Biology and Evolution*, **27**, 921–933.
- Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.
- Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences USA*, **104**, 2785–2790.
- Hibbitts TJ, Fitzgerald LA (2005) Morphological and ecological convergence in two natricine snakes. *Biological Journal of the Linnean Society*, **85**, 363–371.
- Hibbitts TJ, Painter CW, Holycross AT (2009) Ecology of a population of the narrow-headed garter snake (*Thamnophis rufipunctatus*) in New Mexico: catastrophic decline of a river specialist. *The Southwestern Naturalist*, **54**, 461–467.
- Hillis DM, Mable BK, Larson A *et al.* (1996) Nucleic acids IV: sequencing and cloning. In: *Molecular Systematics*, 2nd edn (eds Hillis DM, Moritz C, Mable BK), pp. 342–343. Sinauer, Sunderland, Massachusetts.
- Ho SYM, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Systematic Biology*, **58**, 367–380.
- Holman JA (2000) *Fossil Snakes of North America: Origin, Evolution, Distribution, Paleoecology*. Indiana University Press, Bloomington, IN.
- Holycross AT, Burger WP, Nigro EJ *et al.* (2006) Surveys for *Thamnophis eques* and *Thamnophis rufipunctatus* in the Gila River Watershed of Arizona and New Mexico. Report submitted to Arizona Game and Fish Department.

- Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.
- Huelsensbeck JP, Ronquist F (2001) MRBAYES: Bayesian inferences of phylogenetic trees. *Bioinformatics*, **17**, 754–755.
- Jeffreys H (1961) *Theory of Probability*. Oxford University Press, Oxford.
- Jennings RD, Christman BL (2009). Pre-monsoonal and post monsoonal habitat use of the Narrow-headed Gartersnake, *Thamnophis rufipunctatus*, along the Gila River. Report submitted to the New Mexico Department of Game and Fish.
- Knowles LL (2000) Tests of Pleistocene speciation in montane grasshoppers from the sky islands of western North America (Genus *Melanoplus*). *Evolution*, **11**, 2623–2635.
- Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623–2635.
- Kubatko LS, Degnan JH (2007) Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, **56**, 17–24.
- Leaché AD (2009) Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). *Systematic Biology*, **58**, 547–559.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- McCormack JE, Peterson AT, Bonaccorso E, Smith TB (2008a) Speciation in the highlands of Mexico: genetic and phenotypic divergence in the Mexican jay (*Aphelocoma ultramarina*). *Molecular Ecology*, **17**, 2505–2521.
- McCormack JE, Bowen BS, Smith TB (2008b) Integrating paleoecology and genetic of bird populations in two sky island archipelagos. *BMC Biology*, **6**, 28.
- McCormack JE, Heled J, Delaney KS *et al.* (2010) Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* Jays. *Evolution*, **65**, 184–202.
- Milne I, Wright F, Rowe G *et al.* (2004) TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics*, **20**, 1806–1807.
- Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear-gene trees. *Evolution*, **49**, 718–726.
- Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology*, **3**, 401–411.
- Nielsen R (2005) Molecular signatures of natural selection. *Annual Reviews of Genetics*, **39**, 197–218.
- Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a Markov Chain Monte Carlo approach. *Genetics*, **158**, 885–896.
- Nowak EM (2006) Monitoring surveys and radio-telemetry of Narrow-headed Gartersnakes (*Thamnophis rufipunctatus*) in Oak Creek, Arizona. Report submitted to the Arizona Game and Fish Department.
- Nowak EM, Santana-Bendix MA (2002) Status, distribution, and management recommendations for the narrow-headed garter snake (*Thamnophis rufipunctatus*) in Oak Creek Arizona. Final report submitted to Arizona Game and Fish Department (Heritage grant 199007). US Geological Survey, Colorado Plateau Research Station.
- Nylander JAA, Wilgenbusch JC, Warren DL *et al.* (2008) AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, **24**, 581–583.
- O'Brien SJ, Evermann JF (1988) Interactive influence of infectious disease and genetic diversity in natural populations. *Trends in Ecology and Evolution*, **3**, 254–259.
- Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568–583.
- Pielou EC (1991) *After the Ice Ages: The Return of Life to Glaciated North America*. University of Chicago Press, Chicago, IL.
- Porter SC, Pierce KL, Hamilton TD (1983) Late Wisconsin mountain glaciation in the western United States. In: *Late-Quaternary Environments of the United States: Vol. 1, The Late Pleistocene* (ed. Porter SC), pp. 71–111. University of Minnesota Press, Minneapolis, MN.
- Posada D (2008) Phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.
- Pyron RA, Burbrink FT (2009) Neogene diversification and stability in the snake tribe Lamproleptini (Serpentes: Colubridae). *Molecular Phylogenetics and Evolution*, **52**, 524–529.
- Quattro JM, Vrijenhoek RC (1989) Fitness differences among remnant populations of the endangered Sonoran topminnow. *Science*, **245**, 976–978.
- de Queiroz A, Lawson R (1994) Phylogenetic relationships of the garter snakes based on DNA sequence and allozyme variation. *Biological Journal of the Linnean Society*, **53**, 209–229.
- de Queiroz A, Lawson R (2008) A peninsula as an island: multiple forms of evidence for overwater colonization of Baja California by the gartersnake *Thamnophis validus*. *Biological Journal of the Linnean Society*, **95**, 409–424.
- de Queiroz A, Lawson R, Lemos-Espinal JA (2002) Phylogenetic relationships of North American Garter Snakes (*Thamnophis*) based on four mitochondrial genes: how much DNA sequence is enough? *Molecular Phylogenetics and Evolution*, **22**, 315–329.
- Raftery AE (1996) Hypothesis testing and model selection. In: *Markov Chain Monte Carlo in Practice* (eds Gilks WR, Spiegelhalter DJ, Richardson S), pp. 163–188. Chapman and Hall, London.
- Rosen PC, Schwalbe CR (1988) Status of the Mexican and Narrow-headed gartersnakes (*Thamnophis eques megalops* and *Thamnophis rufipunctatus*) in Arizona. Arizona Game and Fish Department, unpublished report submitted to US Fish and Wildlife Service, Albuquerque, New Mexico.
- Rossman DA (1995) Taxonomic status of the Southern Durango spotted garter snake, *Thamnophis nigroneuchalis*, Thompson, 1957. *Proceedings of the Louisiana Academy of Science*, **58**, 1–10.
- Saccheri I, Kussarin M, Kankare M *et al.* (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491–494.
- Sanders KL, Lee MSY (2007) Evaluating molecular clock calibrations using Bayesian analyses with soft and hard bounds. *Biology Letters*, **3**, 275–279.
- Schaeffel F, de Queiroz A (1990) Alternative mechanisms of enhanced underwater vision in the garter snakes *Thamnophis melanogaster* and *T. couchii*. *Copeia*, **1**, 50–58.
- Smith CI, Farrell BD (2005) Phylogeography of the longhorn cactus beetle *Moneilema appressum* Leconte (Coleoptera: Cerambycidae): was the differentiation of the Madrean sky islands driven by Pleistocene climate changes? *Molecular Ecology*, **14**, 3049–3065.

- Stephens M, Smith N, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.
- Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time markov chain evolutionary models. *Molecular Biology and Evolution*, **18**, 1001–1013.
- Tanner WW (1985) Snakes of western Chihuahua. *Great Basin Naturalist*, **45**, 615–676.
- Tanner WW (1990) *Thamnophis rufipunctatus*. *Catalogue of Amphibians and Reptiles*, **505**, 1–2.
- Tedford RH, Albright III LB, Barnosky AD *et al.* (2004) Mammalian biochronology of the Arikareean through Hemphillian interval (late Oligocene through early Pliocene epochs). In: *Late Cretaceous and Cenozoic Mammals of North America: Biostratigraphy and Geochronology* (ed. Woodburne M.O.), pp. 169–231. Columbia University Press, New York.
- Thompson FG (1957) A new Mexican gartersnake (Genus *Thamnophis*) with notes on related forms. *Occasional Papers of the Museum of Zoology University of Michigan*, **584**, 1–13.
- Van Devender TR (1990) Late quaternary vegetation and climate of the Chihuahuan desert, United States and Mexico. In: *Packrat Middens: The Last 40,000 Years of Biotic Change* (eds Betancourt JL, Van Devender TR, Martin PS), pp. 104–133. University of Arizona Press, Tucson, Arizona.
- Wood DA, Fisher RN, Reeder TW (2008) Novel patterns of historical isolation, dispersal, and secondary contact across Baja California in the Rosy Boa (*Lichanura trivirgata*). *Molecular Phylogenetics and Evolution*, **46**, 484–582.

D.W. is a wildlife geneticist with the U.S. Geological Survey. His research integrates phylogenetic and population genetic approaches to understand processes that contribute to biodiversity. He is also inordinately interested in reptiles, particularly snakes. A.V. is a research geneticist with the U.S. Geological Survey. Her research focuses on landscape genetics

and applications to conservation. J.L. is a research professor of ecology in the Universidad Nacional Autónoma de México (FES Iztacala-UNAM). He is interested in the ecology and distribution of Mexican amphibians and reptiles, especially from the northern portion of the country. A.H. is a professor at Mesa Community College and research professor at Arizona State University. His research integrates natural history, ecological, behavioral, and genetic data and applies it to the conservation and management of threatened populations.

Data accessibility

Please see Appendix I. DNA sequences: GenBank Accessions NADH1 (JF946393–JF946484, JF974262–JF974283); VIM5 (JF946082–JF946184); TATABOX (JF946185–JF946288); R35 (JF946289–JF946392).

Supporting information

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

Fig. S1 Individual Bayesian phylogenetic gene trees for the three nuclear loci (TATA, VIM5, and R35). Individuals sampled from each of the three isolates '*rufipunctatus*', '*unilabialis*', and '*nigromuchalis*' are colored blue, orange and red, respectively. The numbers on branches indicate Bayesian posterior probabilities (Pp).

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

Appendix I

Major drainage, site locality and specimen information for *Thamnophis rufipunctatus* species complex used in this study.

| Major drainages | Site code | Site locality | Latitude | Longitude | Sequence data (GenBank Accession no.) | | | | | | Voucher number | | | | | | | | | | | | |
|-----------------|--|--|------------|------------|---------------------------------------|----------|----------|----------|----------|----------|----------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|--------------|-------------|
| | | | | | NADH1 | TATA | VIM5 | R35 | NADH1 | TATA | | VIM5 | R35 | | | | | | | | | | |
| Salt river | | Mogollon Rim: Arizona | | | | | | | | | | | | | | | | | | | | | |
| | OC1 | Verde River, Oak Creek, Midgley Bridge | 34.88613 | -111.73379 | JF946467 | JF946185 | JF946082 | JF946289 | JF946289 | JF946185 | JF946082 | JF946289 | JF946289 | JF946185 | JF946082 | JF946289 | JF946289 | JF946185 | JF946082 | JF946289 | JF946289 | JF946185 | Tissue only |
| | OC2 | Verde River, Oak Creek, Call of the Canyon | 34.99089 | -111.74444 | JF946468 | JF946186 | JF946083 | JF946290 | JF946290 | JF946186 | JF946083 | JF946290 | JF946290 | JF946186 | JF946083 | JF946290 | JF946290 | JF946186 | JF946083 | JF946290 | JF946290 | JF946186 | Tissue only |
| | OC5 | Verde River, Oak Creek, Manzanita Crossing | 34.93168 | -111.74088 | JF946469 | JF946187 | JF946084 | JF946291 | JF946291 | JF946187 | JF946084 | JF946291 | JF946291 | JF946187 | JF946084 | JF946291 | JF946291 | JF946187 | JF946084 | JF946291 | JF946291 | JF946187 | Tissue only |
| | OC7 | Verde River, Oak Creek, Call of the Canyon | 34.98913 | -111.74487 | JF946470 | JF946188 | JF946085 | JF946292 | JF946292 | JF946188 | JF946085 | JF946292 | JF946292 | JF946188 | JF946085 | JF946292 | JF946292 | JF946188 | JF946085 | JF946292 | JF946292 | JF946188 | Tissue only |
| | OC12 | Verde River, Oak Creek, Midgley Bridge | 34.88440 | -111.74142 | JF946471 | JF946189 | JF946086 | JF946293 | JF946293 | JF946189 | JF946086 | JF946293 | JF946293 | JF946189 | JF946086 | JF946293 | JF946293 | JF946189 | JF946086 | JF946293 | JF946293 | JF946189 | Tissue only |
| | OC15 | Verde River, Oak Creek | 34.98913 | -111.74487 | JF946472 | JF946190 | JF946087 | JF946294 | JF946294 | JF946190 | JF946087 | JF946294 | JF946294 | JF946190 | JF946087 | JF946294 | JF946294 | JF946190 | JF946087 | JF946294 | JF946294 | JF946190 | Tissue only |
| | TC1 | Verde River, Tonto Creek, c. 5 mi S of confluence with Rye Creek | 33.98319 | -111.30048 | JF946473 | JF946191 | JF946088 | JF946295 | JF946295 | JF946191 | JF946088 | JF946295 | JF946295 | JF946191 | JF946088 | JF946295 | JF946295 | JF946191 | JF946088 | JF946295 | JF946295 | JF946191 | ASU35711 |
| | TC2 | Verde River, Haigler Creek, 0.5 km downstream of Alderwood Recreation site | 34.20518 | -110.98694 | JF946474 | JF946192 | JF946089 | JF946296 | JF946296 | JF946192 | JF946089 | JF946296 | JF946296 | JF946192 | JF946089 | JF946296 | JF946296 | JF946192 | JF946089 | JF946296 | JF946296 | JF946192 | Tissue only |
| | BKR5 | Black River, at FR 25 | 33.70340 | -109.45323 | JF946442 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | Tissue only |
| | BKR7 | Black River, at FR 25 | 33.70234 | -109.45418 | JF946443 | JF946193 | JF946090 | JF946297 | JF946297 | JF946193 | JF946090 | JF946297 | JF946297 | JF946193 | JF946090 | JF946297 | JF946297 | JF946193 | JF946090 | JF946297 | JF946297 | JF946193 | Tissue only |
| | BKR9 | Black River, at FR 25 | 33.70324 | -109.45324 | JF946444 | JF946194 | JF946091 | JF946298 | JF946298 | JF946194 | JF946091 | JF946298 | JF946298 | JF946194 | JF946091 | JF946298 | JF946298 | JF946194 | JF946091 | JF946298 | JF946298 | JF946194 | Tissue only |
| | BKR16 | Black River, at FR 25 | 33.70218 | -109.45426 | JF946445 | JF946195 | JF946092 | JF946299 | JF946299 | JF946195 | JF946092 | JF946299 | JF946299 | JF946195 | JF946092 | JF946299 | JF946299 | JF946195 | JF946092 | JF946299 | JF946299 | JF946195 | Tissue only |
| | BKR17 | Black River, at FR 25 | 33.70827 | -109.44656 | JF946446 | JF946196 | JF946093 | JF946300 | JF946300 | JF946196 | JF946093 | JF946300 | JF946300 | JF946196 | JF946093 | JF946300 | JF946300 | JF946196 | JF946093 | JF946300 | JF946300 | JF946196 | Tissue only |
| | BKR18 | Black River, at FR 25 | 33.70219 | -109.45419 | JF946447 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | Tissue only |
| BKR20 | Black River, at FR 25 | 33.70153 | -109.45468 | JF946448 | JF946197 | JF946094 | JF946301 | JF946301 | JF946197 | JF946094 | JF946301 | JF946301 | JF946197 | JF946094 | JF946301 | JF946301 | JF946197 | JF946094 | JF946301 | JF946301 | JF946197 | Tissue only | |
| BKR22 | Black River, at FR 25 | 33.70383 | -109.45300 | JF946449 | JF946198 | JF946095 | JF946302 | JF946302 | JF946198 | JF946095 | JF946302 | JF946302 | JF946198 | JF946095 | JF946302 | JF946302 | JF946198 | JF946095 | JF946302 | JF946302 | JF946198 | Tissue only | |
| BKR23 | Black River, at FR 25 | 33.70140 | -109.45453 | JF946450 | JF946199 | JF946096 | JF946303 | JF946303 | JF946199 | JF946096 | JF946303 | JF946303 | JF946199 | JF946096 | JF946303 | JF946303 | JF946199 | JF946096 | JF946303 | JF946303 | JF946199 | Tissue only | |
| BKR24 | Black River, at FR 25 | 33.70227 | -109.45432 | JF946451 | JF946200 | JF946097 | JF946304 | JF946304 | JF946200 | JF946097 | JF946304 | JF946304 | JF946200 | JF946097 | JF946304 | JF946304 | JF946200 | JF946097 | JF946304 | JF946304 | JF946200 | Tissue only | |
| BKR25 | Black River, at FR 25 | 33.70326 | -109.45329 | JF946452 | JF946201 | JF946098 | JF946305 | JF946305 | JF946201 | JF946098 | JF946305 | JF946305 | JF946201 | JF946098 | JF946305 | JF946305 | JF946201 | JF946098 | JF946305 | JF946305 | JF946201 | Tissue only | |
| BLR1 | Blue River at confluence with Campbell Blue River | 33.71969 | -109.04469 | JF946453 | JF946202 | JF946099 | JF946306 | JF946306 | JF946202 | JF946099 | JF946306 | JF946306 | JF946202 | JF946099 | JF946306 | JF946306 | JF946202 | JF946099 | JF946306 | JF946306 | JF946202 | Tissue only | |
| BLR2 | Blue River at The Box | 33.54639 | -109.19849 | JF946454 | JF946203 | JF946100 | JF946307 | JF946307 | JF946203 | JF946100 | JF946307 | JF946307 | JF946203 | JF946100 | JF946307 | JF946307 | JF946203 | JF946100 | JF946307 | JF946307 | JF946203 | Tissue only | |
| BLR3 | Blue River at The Box | 33.54581 | -109.19949 | JF946455 | JF946204 | JF946101 | JF946308 | JF946308 | JF946204 | JF946101 | JF946308 | JF946308 | JF946204 | JF946101 | JF946308 | JF946308 | JF946204 | JF946101 | JF946308 | JF946308 | JF946204 | Tissue only | |
| BLR4 | Blue River at The Box | 33.54243 | -109.20031 | JF946456 | JF946205 | JF946102 | JF946309 | JF946309 | JF946205 | JF946102 | JF946309 | JF946309 | JF946205 | JF946102 | JF946309 | JF946309 | JF946205 | JF946102 | JF946309 | JF946309 | JF946205 | Tissue only | |
| BLR5 | Blue River at The Box | 33.54412 | -109.19980 | JF946457 | JF946206 | JF946103 | JF946310 | JF946310 | JF946206 | JF946103 | JF946310 | JF946310 | JF946206 | JF946103 | JF946310 | JF946310 | JF946206 | JF946103 | JF946310 | JF946310 | JF946206 | Tissue only | |
| BLR6 | Blue River at The Box | 33.54485 | -109.19965 | JF946458 | JF946207 | JF946104 | JF946311 | JF946311 | JF946207 | JF946104 | JF946311 | JF946311 | JF946207 | JF946104 | JF946311 | JF946311 | JF946207 | JF946104 | JF946311 | JF946311 | JF946207 | Tissue only | |
| BLR7 | Blue River at The Box | 33.54465 | -109.19965 | JF946459 | JF946208 | JF946105 | JF946312 | JF946312 | JF946208 | JF946105 | JF946312 | JF946312 | JF946208 | JF946105 | JF946312 | JF946312 | JF946208 | JF946105 | JF946312 | JF946312 | JF946208 | Tissue only | |
| BLR8 | Blue River, 0.3 mi N lower Juan Miller Road crossing | 33.29715 | -109.19312 | JF946460 | JF946209 | JF946106 | JF946313 | JF946313 | JF946209 | JF946106 | JF946313 | JF946313 | JF946209 | JF946106 | JF946313 | JF946313 | JF946209 | JF946106 | JF946313 | JF946313 | JF946209 | UAZ56664-PSV | |
| | | Mogollon Rim: New Mexico | | | | | | | | | | | | | | | | | | | | | |

Appendix I (Continued).

| Major drainages | Site code | Site locality | Latitude | Longitude | Sequence data (GenBank Accession no.) | | | | | Voucher number | |
|--------------------------------|-----------|---|------------------------------------|------------|---------------------------------------|----------|----------|----------|----------|----------------|-------------|
| | | | | | NADH1 | TATA | VIM5 | R35 | | | |
| Gila River | WC1 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946475 | JF946215 | JF946112 | JF946317 | | Tissue only | |
| | WC3 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | — | — | — | — | | Tissue only | |
| | WC4 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946476 | — | — | — | | Tissue only | |
| | WC5 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946477 | JF946216 | JF946113 | JF946318 | | Tissue only | |
| | WC6 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946478 | JF946217 | JF946114 | JF946319 | | Tissue only | |
| | WC7 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946479 | JF946218 | JF946115 | JF946320 | | Tissue only | |
| | WC8 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946480 | JF946219 | JF946116 | JF946321 | | Tissue only | |
| | WC9 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946481 | JF946220 | JF946117 | JF946322 | | Tissue only | |
| | WC10 | San Francisco River, Whitewater Creek | 33.37332 | -108.84104 | JF946482 | JF946221 | JF946118 | JF946323 | | Tissue only | |
| | WC11 | San Francisco River, Whitewater Creek | 33.37268 | -108.84192 | JF946483 | JF946222 | JF946119 | JF946324 | | Tissue only | |
| | NC1 | San Francisco River, Negro Creek | 33.53827 | -108.60770 | JF946466 | JF946214 | JF946111 | JF946316 | | Tissue only | |
| | SF1 | San Francisco River, at San Francisco Hot Springs | 33.24451 | -108.88173 | AF420174 | — | — | — | | CU12457 | |
| | HB1 | HB1 | Heart Bar Ranch, West Fork Gila R. | 33.20389 | -108.22006 | JF946464 | JF946212 | JF946109 | JF946315 | | Tissue only |
| | | HB2 | Heart Bar Ranch, West Fork Gila R. | 33.20389 | -108.22006 | JF946465 | JF946213 | JF946110 | — | | Tissue only |
| | | GR1 | Gila River Bird Area | 32.84142 | -108.60016 | JF946461 | JF946210 | JF946107 | — | | Tissue only |
| | GR2 | GR2 | Gila River Bird Area | 32.84473 | -108.59474 | JF946462 | — | — | — | | Tissue only |
| | | GR3 | Gila River Bird Area | 32.84473 | -108.59474 | JF946463 | JF946211 | JF946108 | JF946314 | | Tissue only |
| Río Casas Grandes Río Yaqui | LES1 | Mexico: Chihuahua, Sierra Madre Occidental | 29.61485 | -107.87687 | JF946427 | JF946223 | JF946120 | JF946325 | | ESP8874 | |
| | YEP1 | Arroyo La Estancia, c. 6 mi WSW Ignacio Zaragoza | 29.06832 | -107.86386 | JF946440 | JF946228 | JF946125 | JF946330 | | TCWC 92250 | |
| | YEP2 | Arroyo Rincon at HWY 11, c. 1.5 mi N Yepómera | 29.06815 | -107.86366 | JF946441 | JF946229 | JF946126 | JF946331 | | ESP8681 | |
| | TEM1 | South of Temósachic | 28.95244 | -107.83184 | JF946435 | JF946230 | JF946127 | JF946332 | | SDSNH 75636 | |
| | TEM2 | South of Temósachic | 28.95244 | -107.83184 | JF946436 | — | — | — | | TCWC 92249 | |
| | GUE1 | Near Cuidad Guerrero | 28.56589 | -107.49379 | JF946423 | JF946224 | JF946121 | JF946326 | | LSUMZ H8312 | |
| | GUE2 | Near Cuidad Guerrero | 28.56589 | -107.49379 | JF946424 | JF946225 | JF946122 | JF946327 | | LSUMZ H8321 | |
| | GUE3 | Near Cuidad Guerrero | 28.56589 | -107.49379 | JF946425 | JF946226 | JF946123 | JF946328 | | LSUMZ H4297 | |
| | GUE4 | Near Cuidad Guerrero | 28.56589 | -107.49379 | JF946426 | JF946227 | JF946124 | JF946329 | | LSUMZ H8315 | |
| | RSP1 | Off Hwy 16 at Rancho San Pedro, S of Miñaca | 28.39629 | -107.43546 | JF946432 | JF946238 | JF946128 | JF946333 | | SDSNH 75637 | |
| | RSP2 | Off Hwy 16 at Rancho San Pedro, S of Miñaca | 28.39629 | -107.43546 | JF946433 | JF946239 | JF946129 | JF946334 | | SDSNH 75638 | |
| | RSP3 | Off Hwy 16 at Rancho San Pedro, S of Miñaca | 28.39629 | -107.43546 | JF946434 | JF946240 | JF946130 | JF946335 | | TCWC 92245 | |

Appendix I (Continued).

| Major drainages | Site code | Site locality | Latitude | Longitude | Sequence data (GenBank Accession no.) | | | | | Voucher number |
|-----------------|--|--|--|--|--|---|--|---|--|----------------|
| | | | | | NADH1 | TATA | VIM5 | R35 | R35 | |
| TOM1 | Rio Tomochi, Rancho Pinito, c. 7.9 km W Tomochi airport | 28.35838 | -107.93848 | JF946437 | JF946231 | JF946131 | JF946336 | ESP8603 | | |
| TOM2 | Rio Tomochi, Rancho Pinito, c. 7.9 km W Tomochi airport | 28.35838 | -107.93848 | JF946438 | JF946232 | JF946132 | JF946337 | ESP8604 | | |
| TOM3 | Puente Banderillas on Hwy 16, 13.7 mi W Tomochi | 28.34576 | -107.99078 | JF946439 | JF946233 | JF946133 | JF946338 | ESP8888 | | |
| Rio Conchos | BOC1 BOC2 BOC3 BAL1 BAL2 BAL3 CUS1 CUS2 MAG1 MAG2 MAG3 MAG4 VER1 VER2 VER3 VER4 | At Hwy 127, c. 0.5 mi S of Bocoyna At Hwy 127, c. 0.5 mi S of Bocoyna At Hwy 127, c. 0.5 mi S of Bocoyna Rio Balleza at Hwy 24 crossing Rio Balleza at Hwy 24 crossing Rio Balleza at Hwy 24 crossing Arroyo Cusarare, E of Cusarare Arroyo Cusarare, E of Cusarare 12 km N, 11 km E of Maguarichic 12 km N, 11 km E of Maguarichic c. 12 km N, 11 km E of Maguarichic c. 12 km N, 11 km E of Maguarichic Rio Verde at Hwy 24 crossing Rio Verde at Hwy 24 crossing Rio Verde at Hwy 24 crossing Rio Verde at Hwy 24 crossing | 27.83811 27.83811 27.83811 26.65894 26.66106 26.65993 27.59317 27.59317 27.93196 27.93196 27.93196 27.93196 26.28050 26.27838 26.28050 26.28389 | -107.59240 -107.59240 -107.59240 -106.22184 -106.22082 -106.22096 -107.52368 -107.52368 -107.90135 -107.90135 -107.90135 -107.90135 -106.48850 -106.49042 -106.48850 -106.48776 | JF946418 JF946419 JF946420 JF946415 JF946416 JF946417 JF946421 JF946422 JF946428 JF946429 JF946430 JF946431 JF946402 JF946403 JF946404 JF946405 | — JF946241 JF946242 JF946245 JF946246 JF946247 JF946243 JF946244 JF946234 JF946235 JF946236 JF946237 JF946249 JF946250 JF946251 JF946252 | — JF946138 JF946139 JF946142 JF946143 JF946144 JF946140 JF946141 JF946134 JF946135 JF946136 JF946137 — JF946146 JF946147 JF946148 | — JF946343 JF946344 JF946347 JF946348 JF946349 JF946345 JF946346 JF946339 JF946340 JF946341 JF946342 JF946351 JF946352 JF946353 JF946354 | SDSNH 75639 TCWC92247 TCWC92248 Tissue only Tissue only TCWC92244 TCWC92246 SDSNH 75640 ESP8661 ESP8662 ESP8663 ESP8664 Tissue only Tissue only SDSNH 75642 TCWC92243 | |

Appendix I (Continued).

| Major drainages | Site code | Site locality | Latitude | Longitude | Sequence data (GenBank Accession no.) | | | | | Voucher number |
|------------------|-----------|--|----------|------------|---------------------------------------|----------|----------|----------|----------|----------------|
| | | | | | NADH1 | TATA | VIM5 | R35 | R35 | |
| Rio San Lorenzo | VUE1 | Rio Los Vueltes tributary at San Isidro de Calabasas | 25.21362 | -105.96603 | JF946406 | — | — | JF946359 | JF946359 | Tissue only |
| | VUE2 | Rio Los Vueltes tributary at San Isidro de Calabasas | 25.21638 | -105.96595 | JF946407 | JF946257 | JF946153 | JF946360 | JF946360 | TCWC92241 |
| | VUE3 | Rio Los Vueltes tributary at San Isidro de Calabasas | 25.21638 | -105.96595 | JF946408 | JF946258 | JF946154 | JF946361 | JF946361 | Tissue only |
| | VUE4 | Rio Los Vueltes tributary at San Isidro de Calabasas | 25.21044 | -105.96645 | JF946409 | JF946259 | JF946155 | JF946362 | JF946362 | SDSNH 75643 |
| | NIC1 | Rio San Nicolas c. 5 km W of Allares | 25.02838 | -105.94573 | JF946414 | JF946253 | JF946149 | JF946355 | JF946355 | SDSNH 75644 |
| | QUE1 | Arroyo Quebrada La Vueltes at confluence of Hwy 36 | 25.04781 | -105.77134 | JF946410 | JF946254 | JF946150 | JF946356 | JF946356 | Tissue only |
| | QUE2 | Arroyo Quebrada La Vueltes at confluence of Hwy 36 | 25.04781 | -105.77134 | JF946411 | JF946255 | JF946151 | JF946357 | JF946357 | Tissue only |
| | QUE3 | Arroyo Quebrada La Vueltes at confluence of Hwy 36 | 25.04781 | -105.77134 | JF946412 | JF946256 | JF946152 | JF946358 | JF946358 | TCWC92242 |
| Rio Nazas | NAZI | Rio Nazas, at San Rafael de Jicarica | 25.38579 | -104.76962 | JF946413 | JF946248 | JF946145 | JF946350 | JF946350 | TCWC92240 |
| Rio del Presidio | AGU1 | Arroyo del Agua at Hwy 40 crossing | 23.76124 | -105.42358 | JF946393 | JF946261 | JF946157 | JF946364 | JF946364 | Tissue only |
| | AGU2 | Arroyo del Agua at Hwy 40 crossing | 23.76124 | -105.42358 | JF946394 | JF946262 | JF946158 | JF946365 | JF946365 | Tissue only |
| | AGU3 | Arroyo del Agua at Hwy 40 crossing | 23.76124 | -105.42358 | JF946395 | JF946263 | JF946159 | JF946366 | JF946366 | SDSNH 75645 |
| | AGU4 | Arroyo del Agua at Hwy 40 crossing | 23.76124 | -105.42358 | JF946396 | JF946264 | JF946160 | JF946367 | JF946367 | TCWC92239 |
| | AGU5 | 8.5 mi WSW El Salto | 23.73149 | -105.46258 | JF946397 | JF946265 | JF946161 | JF946368 | JF946368 | LSUMZ H4319 |
| | AGU6 | 8.5 mi WSW El Salto | 23.73149 | -105.46258 | JF946398 | — | — | — | — | LSUMZ H4320 |
| | AGU7 | 8.5 mi WSW El Salto | 23.73149 | -105.46258 | JF946399 | JF946268 | — | — | — | LSUMZ H4321 |
| | AGU8 | 5.6 mi WSW El Salto | 23.75882 | -105.42358 | JF946400 | JF946266 | JF946162 | JF946369 | JF946369 | LSUMZ H8276 |
| | AGU9 | 5.6 mi WSW El Salto | 23.75882 | -105.42358 | JF946401 | JF946267 | JF946163 | JF946370 | JF946370 | LSUMZ H8278 |
| | AGU10 | 8.5 mi WSW El Salto | 23.73149 | -105.46258 | JF946484 | JF946260 | JF946156 | JF946363 | JF946363 | LSUMZ H4322 |
| Outgroup Taxa | THME | <i>Thamnophis melanogaster</i> | | | JF974283 | JF946276 | JF946176 | JF946372 | JF946372 | CAS165420 |
| | THVA | <i>Thamnophis validus</i> | | | JF974262 | JF946278 | JF946172 | JF946371 | JF946371 | MVZ236400 |
| | THSC | <i>Thamnophis scaliger</i> | | | JF974263 | JF946271 | JF946167 | JF946373 | JF946373 | CAS214293 |
| | THHA_1 | <i>Thamnophis hammondi</i> | | | JF974281 | JF946272 | JF946170 | JF946377 | JF946377 | MW02-195129 |
| | THHA_2 | <i>T. hammondi</i> | | | JF974282 | JF946273 | JF946171 | JF946377 | JF946377 | TJE14-3A |
| | THEQ_1 | <i>Thamnophis eques</i> | | | JF974279 | JF946279 | JF946173 | JF946375 | JF946375 | ATH 727 |
| | THEQ_2 | <i>T. eques</i> | | | JF974280 | JF946280 | JF946174 | JF946376 | JF946376 | TRJ 965 |

Appendix I (Continued).

| Major drainages | Site code | Site locality | Latitude | Longitude | Sequence data (GenBank Accession no.) | | | | | Voucher number |
|-----------------|----------------------------------|---------------|----------|-----------|---------------------------------------|----------|----------|------------|--|----------------|
| | | | | | NADHI | TATA | VIM5 | R35 | | |
| THCY | <i>Thamnophis cyrtopsis</i> | | | JF974272 | JF946275 | JF946168 | JF946374 | MVZ226240 | | |
| THPU | <i>Thamnophis pulchritatus</i> | | | JF974273 | JF946274 | JF946175 | JF946379 | CAS218285 | | |
| THCH | <i>Thamnophis chrysocephalus</i> | | | JF974274 | JF946277 | JF946169 | JF946378 | MVZ164784 | | |
| THSA | <i>Thamnophis sauritus</i> | | | JF974271 | JF946270 | JF946166 | JF946380 | CAS204801 | | |
| THSL_1 | <i>Thamnophis sirtalis</i> | | | JF974277 | JF946269 | JF946164 | JF946381 | DAW7-2-1 | | |
| THSL_2 | <i>T. sirtalis</i> | | | JF974278 | JF946269 | JF946165 | JF946382 | DAW7-4-1 | | |
| STDE | <i>Storeria dekayi</i> | | | JF974266 | JF946281 | JF946179 | JF946385 | MVZ233662 | | |
| NEFA_1 | <i>Nerodia fasciata</i> | | | JF974275 | JF946283 | JF946178 | JF946383 | RNF8337 | | |
| NEFA_2 | <i>N. fasciata</i> | | | JF974276 | JF946282 | JF946177 | JF946384 | RNF8338 | | |
| NATE_1 | <i>Natrix tessellata</i> | | | JF974264 | JF946285 | JF946184 | JF946387 | CAS 219930 | | |
| NATE_2 | <i>Natrix tessellata</i> | | | JF974265 | JF946286 | — | JF946388 | MVZ 233292 | | |
| NANA_1 | <i>Natrix natrix</i> | | | JF974268 | — | JF946180 | JF946389 | CAS 175879 | | |
| NANA_2 | <i>N. natrix</i> | | | JF974267 | JF946284 | JF946181 | JF946390 | CAS 219929 | | |
| NAMA_1 | <i>Natrix maura</i> | | | JF974269 | JF946288 | JF946182 | JF946391 | MVZ 235726 | | |
| NAMA_2 | <i>N. maura</i> | | | JF974270 | JF946287 | JF946183 | JF946392 | MVZ 186256 | | |

Nucleotide sequence data collected for new individuals are denoted by GenBank Accession nos, and missing data are denoted with ‘—’. Abbreviations and field series abbreviations are as follows: ASU, Arizona State University; ATH, Andrew T. Holycross; CAS, California Academy of Sciences; EMN, Erika M. Nowak; ESP, Charles W. Painter; LSUMZ, Louisiana State University Museum of Zoology; MVZ, Museum of Vertebrate Zoology; RNF, Robert N Fisher; SDSNH, San Diego Natural History Museum; TCWC, Texas Cooperative Wildlife Collection; TRJ, Thomas R. Jones; UAZ, University of Arizona.