Received: 27 April 2021 | Revised: 4 August 2021 | Accepted: 16 August 2021

DOI: 10.1111/mec.16147

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Genomic signatures of inbreeding and mutation load in a threatened rattlesnake

Alexander Ochoa | H. Lisle Gibbs

Department of Evolution, Ecology, and Organismal Biology, Ohio Biodiversity Conservation Partnership, Ohio State University, Columbus, Ohio, USA

Correspondence
Alexander Ochoa, Department of Biology, University of Central Florida, Orlando, FL, USA.
Email: alexander.ochoa@ucf.edu

H. Lisle Gibbs, Department of Evolution, Ecology, and Organismal Biology, Ohio Biodiversity Conservation Partnership, Ohio State University, Columbus, OH, USA.
Email: gibbs.128@osu.edu

Funding information
State Wildlife Grants Program, Ohio Division of Wildlife and US Fish and Wildlife Service; Division of Environmental Biology, Grant/Award Number: 1638872

Abstract

Theory predicts that threatened species living in small populations will experience high levels of inbreeding that will increase their genetic load, but recent work suggests that the impact of load may be minimized by purging resulting from long-term population bottlenecks. Empirical studies that examine this idea using genome-wide estimates of inbreeding and genetic load in threatened species are limited. Here we use individual genome resequencing data to compare levels of inbreeding, levels of genetic load (estimated as mutation load) and population history in threatened Eastern massasauga rattlesnakes (*Sistrurus catenatus*), which exist in small isolated populations, and closely related yet outbred Western massasauga rattlesnakes (*Sistrurus tergeminus*). In terms of inbreeding, *S. catenatus* genomes had a greater number of runs of homozygosity of varying sizes, indicating sustained inbreeding through repeated bottlenecks when compared to *S. tergeminus*. At the species level, outbred *S. tergeminus* had higher genome-wide levels of mutation load in the form of greater numbers of derived deleterious mutations compared to *S. catenatus*, presumably due to long-term purging of deleterious mutations in *S. catenatus*. In contrast, mutations that escaped species-level drift effects within *S. catenatus* populations were in general more frequent and more often found in homozygous genotypes than in *S. tergeminus*, suggesting a reduced efficiency of purifying selection in smaller *S. catenatus* populations for most mutations. Our results support an emerging idea that the historical demography of a threatened species has a significant impact on the type of genetic load present, which impacts implementation of conservation actions such as genetic rescue.

KEYWORDS
demographic bottlenecks, inbreeding, mutation load, ROHs, *Sistrurus catenatus*

1 | INTRODUCTION

Threatened and endangered species often exist in small isolated populations that theory predicts may suffer elevated risks of extinction due to impacts on their genetic makeup or through demographic stochasticity (Caughley, 1994; Frankham et al., 2017; van Oosterhout, 2020). Populations with these characteristics experience increased genetic drift and inbreeding that lead to a loss of genetic variation, a reduction in the efficacy of natural selection, and an increase in the frequency and expression of deleterious recessive mutations termed genetic load (Charlesworth, 2009; Keller & Waller, 2002). However, the precise genetic endpoint of this process in terms of observed levels of genetic load in natural populations depends on complex interactions between the levels of inbreeding and the effectiveness of purifying selection as mediated through population size (Grossen et al., 2020; Hedrick & Garcia-Dorado, 2016).

Specifically, inbreeding results in the increased expression of recessive deleterious mutations that create the potential for selection

to reduce the frequency of these mutations, depending on the degree of dominance and the magnitude of the deleterious effects (Glémin, 2003). This process, termed genetic purging, is most effective at high levels of inbreeding, and so bottlenecks tend to lead to selection purging highly deleterious, recessive mutations unless population sizes are extremely small (García-Dorado, 2012; Kirkpatrick & Jarne, 2000). However, reduced population size also increases genetic drift and reduces the efficacy of selection, allowing deleterious mutations to drift to higher frequencies (Renaut & Rieseberg, 2015; Robinson et al., 2016). Thus, the type and level of genetic load present in a given population reflect a balance of the relative magnitudes of these distinct processes (inbreeding vs. reduced selection due to drift) with opposite effects on mutation frequencies. As a result, assessing levels of inbreeding, effective population size (N_e) and levels of genetic load is an important goal of empirical studies that seek to evaluate both the genetic risks impacting the long-term viability of threatened species (Benazzo et al., 2017; Grossen et al., 2020; Mathur & DeWoody, 2021; Robinson et al., 2016, 2019) and the appropriateness of specific management activities such as genetic rescue for mitigating these risks (Kyriazis et al., 2020; but see Ralls et al., 2020).

In terms of causes of variation in genetic load at the species level, van der Valk et al. (2021) have recently emphasized the previously underappreciated impact of historical demography on the observed level of genetic load of deleterious mutations in contemporary populations. They have shown that estimates of genetic load vary widely across mammal species with small N_e sizes and are only weakly correlated with genome-wide levels of inbreeding. They argue that this is because long-term reductions in N_e can result in sustained purging of deleterious mutations, leading to reduced genetic load in present-day populations. Such a pattern has been observed in populations of endangered species which have experienced long-term bottlenecks (Benazzo et al., 2017), suggesting the importance of incorporating historical patterns of population size in explanations for genetic load.

Whole genome resequencing offers new data and methods of analysis to evaluate inbreeding and genetic load in endangered species (Brüniche-Olsen et al., 2018). For assessing inbreeding, individual genome sequences provide information on runs of homozygosity (ROHs), which are genomic regions that have identical haplotypes and are identical by state (Brüniche-Olsen et al., 2018; Ceballos et al., 2018). These ROH regions arise in a genome due to the transmission of genome segments that are identical by descent from parents to offspring (Ceballos et al., 2018). ROH segments can be used to quantify inbreeding reflected by genomic autozygosity because ROH distributions more accurately reflect actual levels of inbreeding than do pedigree histories (Kardos et al., 2015).

Distributions of different size ROHs also provide insights into recent demography because under random mating, the length of ROH regions is expected to decrease with increasing number of generations to the most recent common ancestor due to recombination and *de novo* mutations (Bosse et al., 2012). Therefore, the number and length of ROHs reflect the timing and intensity of a population bottleneck, with longer ROHs originating from recent inbreeding, and shorter ROHs from ancestral bottlenecks (Ceballos et al., 2018).

For assessing genetic load, new analytical methods (Adzhubei et al., 2010; Choi et al., 2012) now permit strong inference of the mutational effects of substitutions in coding sequences assumed to be under strong purifying selection. As such, data from these approaches can be used to estimate levels of accumulation of putatively deleterious mutations at the population or species level in nonmodel species and to explore the process that drives variation in genetic load (estimated as mutation load) in natural populations (Allendorf, 2017; Kardos et al., 2016).

Resolving how genetic factors impact the future viability of small populations at the species and population levels is an important goal of management plans for species currently found in small isolated populations (Szymanski et al., 2016). One such species is the Eastern Massasauga rattlesnake (*Sistrurus catenatus*). Historically, this small snake was found in wetlands and surrounding upland habitat in midwestern and eastern North America. Population declines throughout its range due to habitat fragmentation and destruction have led to the listing of this species as threatened under the United States Endangered Species Act (U.S. Fish & Wildlife Service, 2016) and as a Species at Risk in Canada (Government of Canada, 2009). This snake has a polygynous mating system (Jellen et al., 2007), and genetic analysis of parentage has shown direct evidence for multiple paternity in single litters (Stedman et al., 2016). This species exhibits little phylogeographical structure across its range—and hence no evolutionary significant conservation units (Sovic et al., 2016)—and high levels of population genetic structure with little to no migration between populations even at local geographical scales (Chiucchi & Gibbs, 2010; Martin et al., 2021; Sovic et al., 2019). Thus, the relevant management units within this species are individual populations.

Demographic studies suggest that the negative genetic impacts of drift on population viability are potentially a significant conservation issue for *S. catenatus*. In particular, Sovic et al. (2019) have recently shown that the contemporary N_e values of most populations of *S. catenatus* are <50 individuals and that many have undergone recent bottlenecks, suggesting that genetic drift and the negative genetic impacts of inbreeding could be significant yet evidence in support of this possibility is unclear. Specifically, heterozygosity–fitness correlations based on neutral genetic markers and an indirect estimate of fitness (body condition) show few positive relationships consistent with the negative effects of drift and inbreeding, but the power of these tests may be limited because body condition is only an indirect measure of individual fitness (Gibbs & Chiucchi, 2012; Sovic et al., 2019). Genome-level quantification of levels of inbreeding and genetic load would help address the impact of genetic factors on population viability in this species (Benazzo et al., 2017; Grossen et al., 2020; Mathur & DeWoody, 2021; Robinson et al., 2019).

To this end, we used data from resequenced genomes from 90 *S. catenatus* individuals from nine populations to conduct detailed analyses of inbreeding, recent population demography and levels of mutation load at the species and population level. To provide

necessary context for our results, we also generated similar data for 10 individuals from a single population of a closely related yet non-threatened species, the Western Massasauga rattlesnake (*Sistrurus tergeminus*). In contrast to *S. catenatus*, *S. tergeminus* is relatively common throughout much of its range, and genetic studies show no evidence for high levels of population genetic structure, supporting the idea that it represents an outbred species that has experienced far fewer anthropogenic impacts on its genetic makeup (Bylsma, 2020; McCluskey & Bender, 2015; Ryberg et al., 2015).

We used these data to address the following questions: What are the genome-wide levels of inbreeding and is inbreeding higher in *S. catenatus* as compared to *S. tergeminus*? Based on demographic analyses that use ROH size and abundance distributions, what are the trends in population size over recent timescales and how do these aid in interpreting observed levels of inbreeding and levels of mutation load in each species? What are levels of different types of deleterious mutations in both species and among *S. catenatus* populations and what evolutionary processes underlie differences in load between species and *S. catenatus* populations? Our results show how analysis of genomic data can test the idea that the historical demography of a threatened species has a significant impact on the type of genetic load present, which impacts conservation actions such as genetic rescue, and also identify populations of *S. catenatus* whose viability may be at risk due to genetic factors.

2 | MATERIALS AND METHODS

2.1 | Samples and DNA sequencing and processing

We analysed whole-genome sequences from 90 *S. catenatus* individuals from nine populations located in the Great Lakes region of the USA and Canada. We also analysed whole-genome sequences from 10 *S. tergeminus* individuals from a population located in Cheyenne Bottoms, Kansas, as a reference for expected patterns of variation in a closely related, yet outbred species (see above). Locations, population names, codes and sample sizes are presented in Figure 1 and Table S1.

We collected, sequenced and processed DNA samples as described in Section S1. Briefly, we used TRIM GALORE version 0.4.5 5 (https://github.com/FelixKrueger/TrimGalore) to trim sample Illumina paired-end (PE) 150-bp reads. We then used BWA version 0.7.15 (Li & Durbin, 2009) to map the resulting PE reads to an *S. catenatus* reference genome (~1.6 Gb; Broe et al., in prep.). Next, we used a combination of SAMTOOLS version 1.3.1 (Li et al., 2009) and GATK version 3.8 (McKenna et al., 2010) to increase the quality of the mapping. Finally, we used MASHMAP version 2.0 (Jain et al., 2018) and a prairie rattlesnake (*Crotalus viridis*) genome assembly (Pasquesi et al., 2018) to identify and exclude nonautosomal scaffolds from our data. For most analyses, we downsampled data to a mean sequence depth

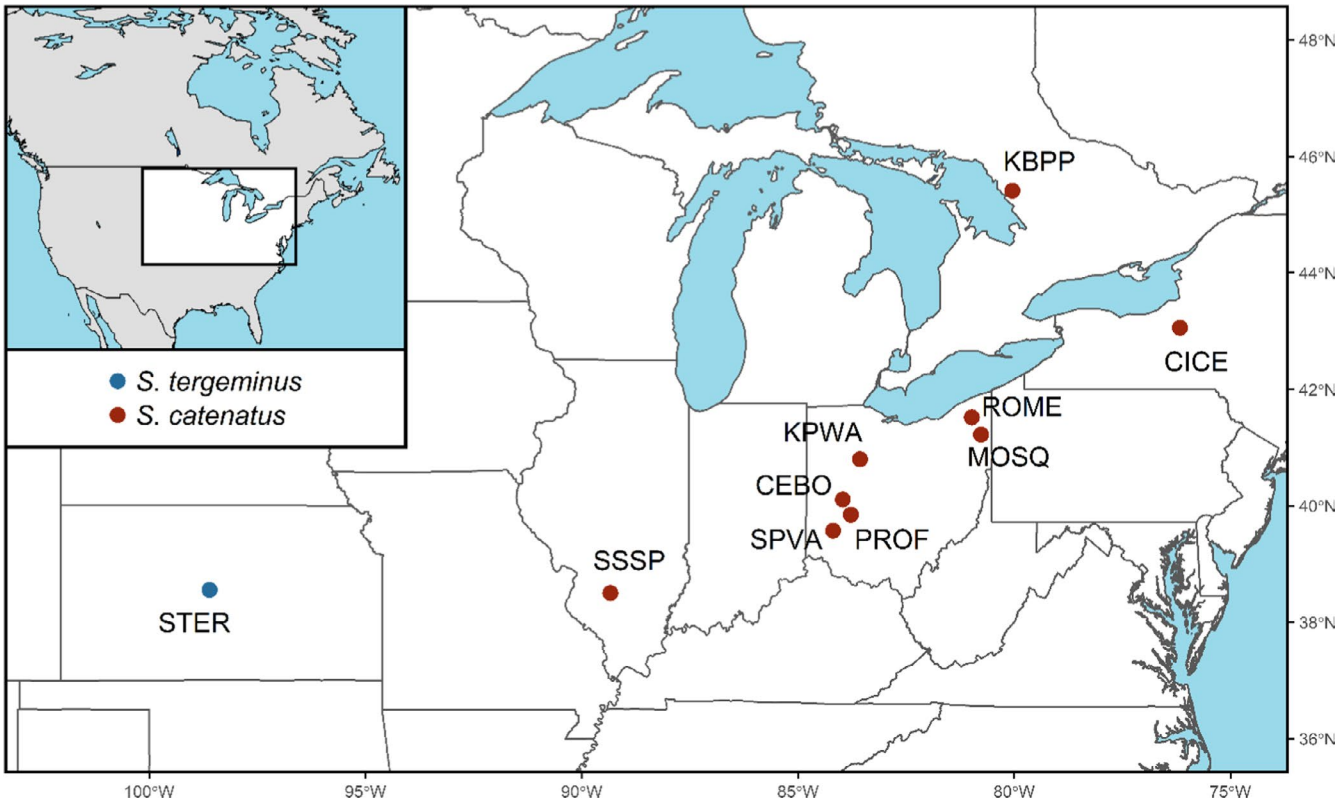

FIGURE 1 Geographical location in the USA and Canada of the only *S. tergeminus* population (blue circle) and the nine *S. catenatus* populations (red circles) analysed in this study. Population names, codes and sample sizes are as follows: Cheyenne Bottoms (STER, $n = 10$), South Shore State Park (SSSP, $n = 11$), Spring Valley (SPVA, $n = 10$), Cedar Bog Nature Preserve (CEBO, $n = 9$), Prairie Road Fen (PROF, $n = 10$), Killdeer Plains Wildlife Area (KPWA, $n = 10$), Rome (ROME, $n = 10$), Mosquito Creek (MOSQ, $n = 10$), Cicero (CICE, $n = 11$) and Killbear Provincial Park (KBPP, $n = 9$) [Colour figure can be viewed at wileyonlinelibrary.com]

of 5× (see Table S1 for mean sequence depths prior to downsampling) unless indicated otherwise.

2.2 | Genetic diversity and inbreeding

We compared measures of genetic diversity, size and abundance of ROHs, and relatedness at the inter- and intraspecific levels using autosomal scaffolds ≥2 Mb (Figure S1), which covered 23% of the *S. catenatus* reference genome. We used ANGSD version 0.930 (Korneliussen et al., 2014) to estimate nucleotide diversity (π; Nei & Li, 1979), heterozygosity at equilibrium (θ_W; Watterson, 1975), and weighted differences between π and θ_W (*D*; Tajima, 1989) for each population based on the site–frequency–spectrum principle (Nielsen et al., 2012). We also calculated long-term effective sizes ($N_{e.LT}$) from the equation $\theta_W = 4N_{e.LT}\mu$ (Watterson, 1975) after assuming a generation time of 3 years (Sovic et al., 2019) and a neutral mutation rate of $\mu = 7.2 \times 10^{-9}$ substitutions per site per generation (see squamates in Green et al., 2014).

For each sample, we used ROHAN version 1.0 (Renaud et al., 2019) to infer the size and distribution of ROH tracts across scaffolds, θ_W outside of ROHs ($\theta_{W.OUT}$), fraction of the genome covered by ROHs (F_{ROH}) and number of ROHs (N_{ROH}). Following Benazzo et al. (2017), we defined ROHs as genomic regions ≥50 kb with a heterozygosity rate $\leq 5 \times 10^{-4}$ (i.e., ≤25 heterozygous genotypes in 50-kb sliding windows), thus accounting for potential sequencing errors. We quantified the impact of inbreeding on levels of variation by comparing individual estimates of F_{ROH} with individual estimates of $\theta_{W.OUT}$ and N_{ROH}. Inbred individuals should show a high F_{ROH} and low $\theta_{W.OUT}$ (Saremi et al., 2019), whereas individuals from populations that have experienced a recent bottleneck should show an excess of N_{ROH} (Ceballos et al., 2018). To conduct cross-study comparisons with other threatened and endangered species and subspecies (Benazzo et al., 2017; Grossen et al., 2020; Robinson et al., 2019; Saremi et al., 2019; van der Valk et al., 2021), we recalculated F_{ROH} for ROH sizes ≥0.1, 1, 2 and 2.5 Mb, which are specific values used in these other studies.

Estimates of individual relatedness can provide an additional evaluation of the level of inbreeding occurring within populations. As such, we used ANGSD and NGSRELATE version 2 (Korneliussen & Moltke, 2015) to calculate relatedness for pairs of individuals from the same population based on genotype likelihood distributions. We used the r_{xy} statistic (Hedrick & Lacy, 2015) because it is designed for estimating individual relatedness in populations where inbreeding occurs.

2.3 | Recent demographic history

To describe recent inter- and intraspecific demographic events based on the ROH-size data generated by ROHAN, we first used MCLUST version 3 in R (R Core Team, 2020) to identify distinct ROH-size clusters (see Pemberton et al., 2012) from a pooled distribution of *S. catenatus* samples. Since Schield et al. (2020) have documented significant differences in recombination rates on snake macro- and microchromosomes that could bias demographic inferences (see below), we excluded microchromosomal scaffolds (Figure S1) from this analysis.

We then dated the resulting ROH-size clusters using the equation $g = 100/(2rL)$, where g is the generation time (i.e., 3 years, as above), r is the recombination rate in centimorgans (cM) Mb^{-1}, and L is the length of each ROH-size cluster inMb (Saremi et al., 2019). Because there are no estimates of r for any reptilian species, we used a recombination rate estimate of 2.8 cM Mb^{-1} for chicken (*Gallus gallus*) macrochromosomes (International Chicken Genome Sequencing Consortium, 2004).

2.4 | Long-term demographic history

We assessed long-term species-specific demographic histories using PSMC version 0.6.4 (Li & Durbin, 2011), which infers changes in N_e over extended periods of time based on local densities of heterozygous genotypes across a single diploid genome. We used samples with the highest mean sequence depth from each species (*S. catenatus*: sca1038, ~29×; *S. tergeminus*: ste0121, ~16×; Table S1) and GATK's best practices pipeline (www.broadinstitute.org) to call single nucleotide polymorphisms (SNPs) across autosomal scaffolds of all lengths after excluding genomic regions containing repetitive elements, as defined by REPEATMASKER version 4.0.7 (Smit et al., 2010). We ran PSMC with these data for 100 bootstrap replicates and assumed a neutral mutation rate of 7.2×10^{-9} substitutions per site per generation, as above.

2.5 | Mutation load

We evaluated mutation load at the species and population levels using variants detected in a set of conserved single-copy coding loci (Benchmarking Universal Single-Copy Orthologs or BUSCOs; Simão et al., 2015) that are expected to be under purifying selection due to evolutionary constraints across specific lineages. To assess mutation load at the species level, we focused on sites containing derived nonsynonymous alleles across species (i.e., species-level mutations); these sites capture differences in genome-wide heterozygosity between species, which can substantially impact estimates of load (van der Valk et al., 2021). To estimate mutation load at the population level, we focused on sites containing derived nonsynonymous alleles within specific populations (i.e., population-level mutations); these sites capture aspects of mutation load that reflect the evolutionary process shaping variation within populations independent of species-level effects.

To examine variation in BUSCO loci, we first performed homology searches for these loci in the *S. catenatus* reference genome. We used BUSCO version 3.1.0 (Simão et al., 2015) and the Tetrapoda version 10 database to identify the genomic coordinates of 5310 query BUSCO loci; we excluded duplicated or fragmented BUSCO loci from further analyses. We then used ANGSD and data from all

autosomal scaffolds to characterize SNPs and genotypes within these loci across samples. To identify derived nonsynonymous alleles, we translated and compared data to a genome assembly of a closely-related outgroup (*Sistrurus miliarius*; Broe et al., in prep.; see Section S2 for detailed bioinformatic methods).

We used PROVEAN version 1.1 (Choi & Chan, 2015) to determine the potential deleterious impact of the derived nonsynonymous alleles in the BUSCO loci (Perrier et al., 2017; Renaut & Rieseberg, 2015). PROVEAN uses BLAST searches of protein databases and subsequent alignments of protein sequences from related species to generate a measure (i.e., PROVEAN score) of how phylogenetically conserved a given derived nonsynonymous allele is among homologous protein sequences. We used these scores to classify derived nonsynonymous alleles into "benign" (PROVEAN scores > −2.5), "mildly deleterious" (PROVEAN scores ≤ −2.5 and >−4.1), and "moderately deleterious" (PROVEAN scores ≤ −4.1). We also identified "highly deleterious" derived nonsynonymous alleles that resulted in intermediate stop codons; since these mutations are not accounted for by PROVEAN, we assigned them a conservative score of −12, which roughly represents the lowest score detected across variants. For each sample, we weighted derived alleles—and genotypes carrying these alleles—with data contained in the corresponding mutation class sites (i.e., benign, mildly deleterious, moderately deleterious, highly deleterious) identified across species and within populations to estimate species- and population-level mutation load, respectively.

We also calculated the *Rxy* statistic for selected pairs of populations. Broadly, this metric reflects the relative frequency of a specific class of deleterious mutations with respect to the frequency of neutral mutations in one population over that of another one. In essence, *Rxy* standardizes for sampling effects and background differences resulting from population substructure (Do et al., 2015; Grossen et al., 2020). We used derived synonymous alleles found in the BUSCO loci as background neutral variation for the analysis. For a species-level assessment of the *Rxy* statistic, we compared population KPWA (*S. catenatus*) with respect to STER (*S. tergeminus*); we selected KPWA for this analysis because multiple lines of evidence suggest that it is one of the largest *S. catenatus* populations in our sample (Sovic et al., 2019). Within *S. catenatus*, we compared PROF with respect to KPWA, as the former is one of the smallest populations in our sample (Sovic et al., 2019). We generated standard errors (*SE*) for *Rxy* estimates by jackknifing *Rxy* values associated with individual scaffolds from each pairwise comparison.

Throughout our study we used nonparametric Wilcoxon–Mann–Whitney tests for statistical comparisons between groups. We further used the false discovery rate method (Benjamini & Hochberg, 1995) for *p*-value adjustments across multiple tests.

3 | RESULTS

3.1 | Genetic diversity and inbreeding

Table 1 shows different measures of genetic diversity for multiple populations of *S. catenatus* and the single reference population of *S. tergeminus*. The general patterns match those found with previous analyses of much smaller genetic data sets from these species (McCluskey & Bender, 2015; Sovic et al., 2019). Overall, *S. catenatus* shows substantially less genetic diversity and much smaller $N_{e.LT}$ sizes than *S. tergeminus*, while genetic diversity and $N_{e.LT}$ sizes across individual populations of *S. catenatus* are variable.

In terms of inbreeding, fine-scale representations of the distribution of ROHs across the longest scaffold analysed suggest that such tracts are longer and more abundant in threatened *S. catenatus* individuals relative to nonthreatened *S. tergeminus* individuals (Figure 2). Further analyses confirm that *S. catenatus* shows low genetic diversity and high inbreeding levels with respect to *S. tergeminus* at the genome-wide level (Figure 3). Specifically, when averaged across all individuals, *S. catenatus* has a mean genome-wide $\theta_{W.OUT}$ (scaled to 10^3 bp hereafter) of 2.62 (*SE* = 0.01) and F_{ROH} of 0.551 (*SE* = 0.013), while *S. tergeminus* has a mean $\theta_{W.OUT}$ of 4.41 (*SE* = 0.02) and F_{ROH} of

TABLE 1 Genetic diversity and effective size summary statistics for *S. tergeminus* and *S. catenatus* populations based on whole-genome data

Species	Population	$\pi \pm SE$ ($\times 10^{-3}$)	$\theta_W \pm SE$ ($\times 10^{-3}$)	$D \pm SE$	$N_{e.LT} \pm SE$	$N_{e.ST}$ (95% CI)
S. tergeminus	STER	8.73 ± 0.09	7.51 ± 0.07	0.68 ± 0.01	260,641 ± 2600	NA
S. catenatus	SSSP	2.38 ± 0.05	2.18 ± 0.04	0.34 ± 0.03	75,571 ± 1307	22 (9–Inf)
	SPVA	2.36 ± 0.05	2.27 ± 0.04	0.11 ± 0.03	78,909 ± 1254	8 (3–19)
	CEBO	2.09 ± 0.04	1.90 ± 0.03	0.40 ± 0.03	65,823 ± 1048	NA
	PROF	2.34 ± 0.05	2.34 ± 0.03	−0.04 ± 0.03	81,191 ± 1185	24 (18–32)
	KPWA	2.35 ± 0.05	2.18 ± 0.04	0.29 ± 0.03	75,655 ± 1326	44 (37–53)
	ROME	2.14 ± 0.04	2.18 ± 0.03	−0.12 ± 0.03	75,700 ± 1026	3 (3–3)
	MOSQ	1.93 ± 0.04	1.65 ± 0.03	0.68 ± 0.04	57,120 ± 942	7 (6–8)
	CICE	2.41 ± 0.05	2.39 ± 0.03	0.00 ± 0.03	83,064 ± 1143	13 (10–16)
	KBPP	2.09 ± 0.04	1.82 ± 0.03	0.58 ± 0.03	63,277 ± 1002	10 (9–12)

Estimates are given for the following parameters: nucleotide diversity (π), heterozygosity at equilibrium (θ_W), weighted differences between π and θ_W (*D*), long-term effective sizes ($N_{e.LT}$) and short-term effective sizes ($N_{e.ST}$). $N_{e.ST}$ values for *S. catenatus* populations were obtained from Sovic et al. (2019). Standard errors (*SE*) or 95% confidence intervals (CI) for each parameter are also indicated.

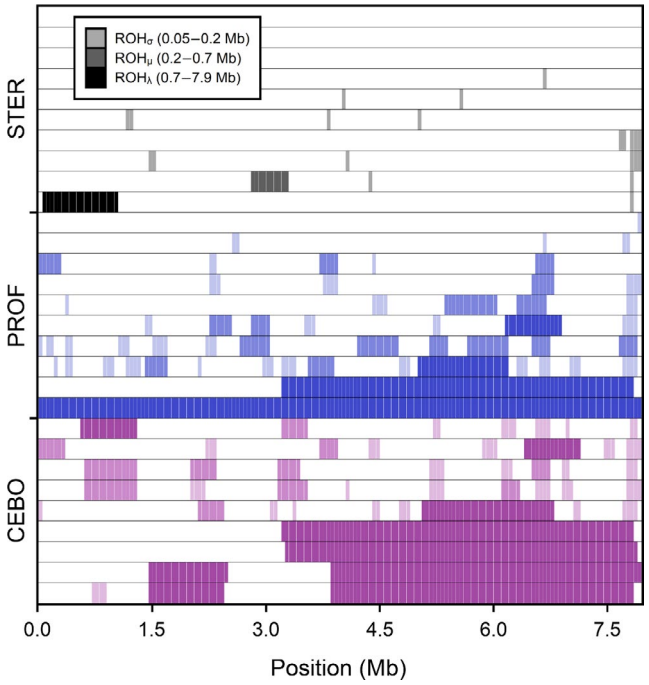

FIGURE 2 Distribution of short (ROH_{σ} = 0.05–0.2 Mb), medium (ROH_{μ} = 0.2–0.7 Mb) and long (ROH_{λ} = 0.7–7.9 Mb) runs of homozygosity along the longest scaffold analysed (~7.9 Mb) for individuals (horizontal bars) from the reference *S. tergeminus* population (STER—black) and two *S. catenatus* populations (PROF—blue; CEBO—purple) with low and high levels of inbreeding. ROH size classes are indicated by gradual color shadings (see box legend) for each population [Colour figure can be viewed at wileyonlinelibrary.com]

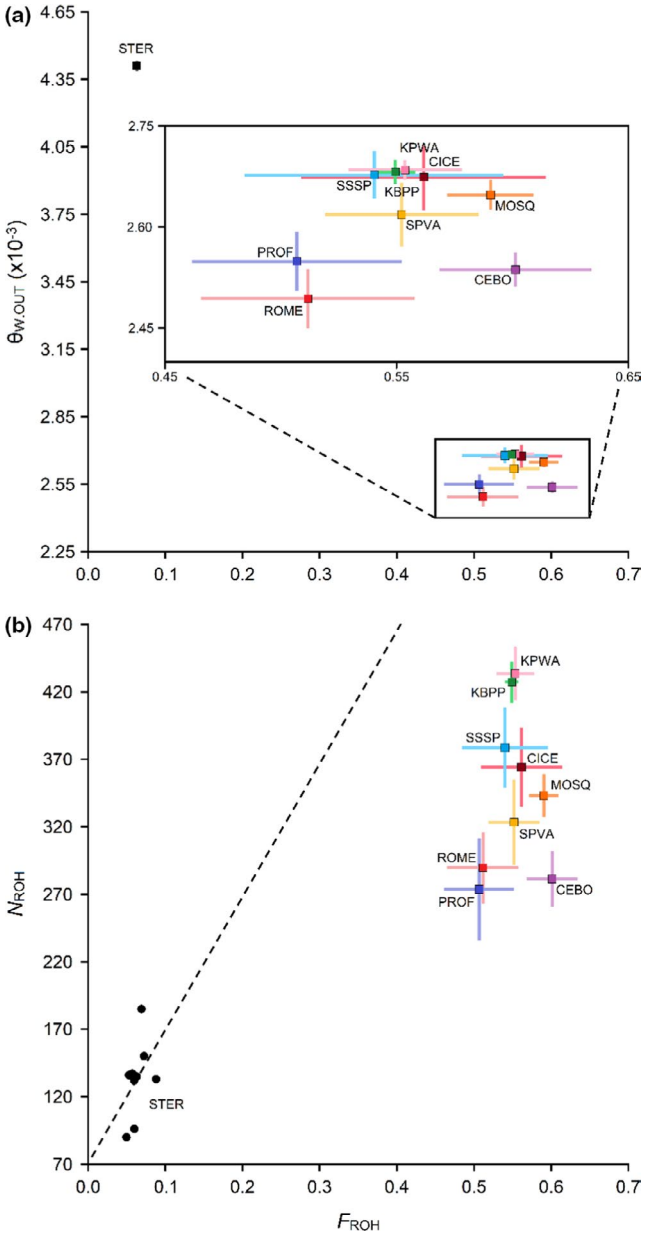

FIGURE 3 Scatter plots of (a) $\theta_{W.OUT}$ and (b) N_{ROH} with respect to F_{ROH} for *S. tergeminus* (STER) and *S. catenatus* (SSSP, SPVA, CEBO, PROF, KPWA, ROME, MOSQ, CICE, KBPP) populations. Vertical and horizontal bars reflect standard errors from the estimated mean (squares) across individuals from the same population. For the plot in (b), each STER individual is represented by a black circle; the dotted line represents the association between N_{ROH} and F_{ROH} for these individuals after performing a linear regression analysis [Colour figure can be viewed at wileyonlinelibrary.com]

0.063 (*SE* = 0.004; Figure 3a); differences between species for both metrics are significant ($p < .001$). These patterns are consistent with repeated reductions in N_e and higher levels of inbreeding in *S. catenatus* than in *S. tergeminus*.

A comparison of adjusted F_{ROH} values across studies suggests that *S. catenatus* is moderately inbred with respect to other threatened and endangered species and subspecies (Table S2; also see Brüniche-Olsen et al., 2018). For this analysis, we recalculated F_{ROH} using data from our study to be comparable with the different size criteria used in other studies (see Table S2). We also recognize that such comparisons are problematic because of varying criteria used to define ROH tracts (e.g., maximum number of heterozygous genotypes allowed to account for sequencing errors). However, based on these adjusted comparisons, *S. catenatus* is half as inbred as a population of ~50 Apennine brown bears (*Ursus arctos marsicanus*) that have existed in isolation for ~50 generations (Benazzo et al., 2017), but shows levels of inbreeding similar to those of Alpine ibex (*Capra ibex*), which have recently undergone serial population bottlenecks over the past 8–10 generations (Grossen et al., 2020; Table S2). Our interpretation that *S. catenatus* is moderately inbred is also consistent with F_{ROH} values generated from 100 mammals representing species with varying population sizes (van de Valk et al., 2021). Based on a size criterion of ROH ≥0.1 Mb, the estimated F_{ROH} value for *S. catenatus* (0.546) is in the upper third of values for species and subspecies analysed. On the other hand, F_{ROH} values for *S. tergeminus* are representative of outbred species (Table S2), such as ibex with large, well-connected populations (Grossen et al., 2020).

Pairwise comparisons of $\theta_{W.OUT}$ and F_{ROH} across the nine *S. catenatus* populations examined (Figures 3a and S2) reveal that 28% of these comparisons are significant ($p < .05$) for the $\theta_{W.OUT}$ statistic, whereas no pairwise comparison is significant ($p > .05$ in all cases)

for the F_{ROH} statistic. Values of $\theta_{W.OUT}$ range from 2.49 ($SE = 0.04$) in ROME to 2.69 ($SE = 0.02$) in KPWA, while values of F_{ROH} range from 0.507 ($SE = 0.045$) in PROF to 0.601 ($SE = 0.033$) in CEBO (Figures 3a and S2). These results confirm that genetic diversity (measured as $\theta_{W.OUT}$ in this case) is heterogeneous among *S. catenatus* populations, and suggest that extreme variation in F_{ROH} values within some *S. catenatus* populations (see below) reduces our power to detect statistically significant pairwise comparisons between populations using this metric.

S. catenatus genomes are also composed of a substantially greater number of ROHs ($N_{ROH} = 346$, $SE = 10$) relative to *S. tergeminus* genomes ($N_{ROH} = 133$, $SE = 8$; $p < .001$; Figure 3b). This 2.6-fold increase in N_{ROH} (combined with the 8.7-fold increase in F_{ROH}, as above) is consistent with the effects of recent bottlenecks leading to inbreeding in *S. catenatus* populations, as it implies that ROH tracts have increased in mean length (Figure 3b; also see figure 1 in Ceballos et al., 2018).

We also found substantial variation among individuals for both F_{ROH} and N_{ROH} metrics. For example, there is a 3.8-fold difference in F_{ROH} and 2.7-fold difference in N_{ROH} among individuals from the PROF population (Figure S3). Some of this variation could be due to differences in the degree of relatedness among individuals within single populations, since, unlike the only *S. tergeminus* population analysed, *S. catenatus* populations presented small but perhaps significant (mean = 0.083, $SE = 0.012$) proportions of close relatives ($r_{xy} > .125$; Figure 4). These patterns are also consistent with a recent onset of inbreeding in *S. catenatus* populations.

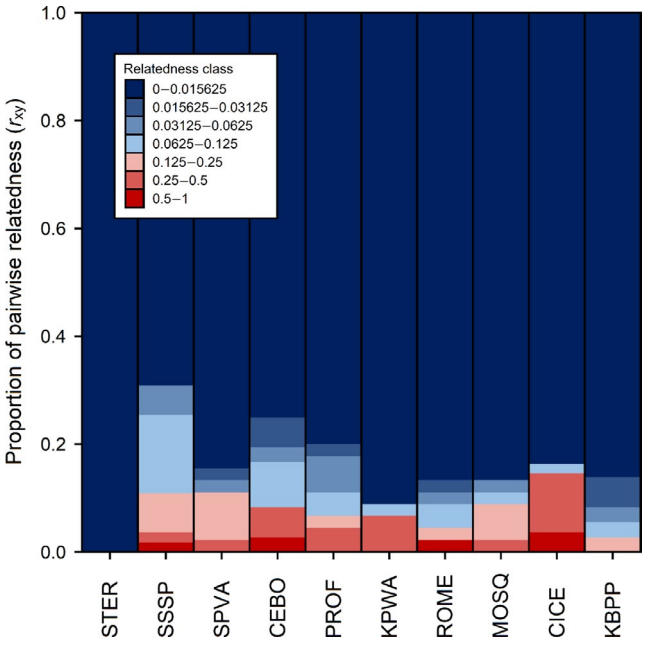

FIGURE 4 Relatedness values (r_{xy}) for pairs of individuals within populations of *S. tergeminus* (STER) and *S. catenatus* (SSSP, SPVA, CEBO, PROF, KPWA, ROME, MOSQ, CICE, KBPP). Colours inside each vertical bar represent the proportion of all pairwise comparisons within a single population that correspond to a particular relatedness class (see box legend) [Colour figure can be viewed at wileyonlinelibrary.com]

3.2 | Recent demographic history

The presence of distinct ROH-size clusters represents timestamps for when population bottlenecks have occurred in the recent past (Bosse et al., 2012; Pemberton et al., 2012). An analysis of the overall distribution of ROH lengths in *S. catenatus* shows evidence for three size classes: short (0.05–0.2 Mb), medium (0.2–0.7 Mb) and long (0.7–7.9 Mb; Figure S4). Assuming a recombination rate of 2.8 cM Mb^{-1} and a generation time of 3 years (see Section 2), we estimated the following time frames for the demographic events responsible for each ROH-size class: short (268–1071 years before present [ybp]), medium (77–268 ybp), and long (7–77 ybp). The last two time frames are consistent with possible anthropogenic impacts on this species (Sovic et al., 2019).

The effects of these events can be assessed by comparing differences in class N_{ROH} between species. As such, N_{ROH} for the three ROH-size classes (i.e., short, medium, long) is always greater in *S. catenatus* relative to *S. tergeminus* (Figure 5a), which is consistent with a greater impact of bottlenecks in *S. catenatus* during the recent past. Specifically, *S. catenatus* genomes have a greater number of short ($N_{ROH} = 172$, $SE = 7$), medium ($N_{ROH} = 100$, $SE = 4$) and long ($N_{ROH} = 60$, $SE = 2$) ROH tracts than *S. tergeminus* genomes (N_{ROH} [short] = 105, $SE = 8$; N_{ROH} [medium] = 15, $SE = 1$; N_{ROH} [long] = 5, $SE < 1$; $p < .01$ in all cases; Figure 5a). Substantial increases in each class suggest that, in *S. catenatus*, ROH-tract gains have been most substantial lately (i.e., 7–77 ybp), followed by ROH-tract gains 77–268 and 268–1071 ybp (Figure 5a).

We can also use the relative proportion of each size class to measure the relative impact of each event in a way that accounts for overall differences in inbreeding between the species (Figure 5c). In this respect, short ROHs account for a 0.505 ($SE = 0.010$) proportion of all ROH tracts in *S. catenatus* and for a 0.831 ($SE = 0.016$) proportion of all ROH tracts in *S. tergeminus* ($p < .001$), whereas the relative proportion of medium (*S. catenatus*: 0.300, $SE = 0.005$; *S. tergeminus*: 0.123, $SE = 0.012$; $p < .001$) and long (*S. catenatus*: 0.196, $SE = 0.009$; *S. tergeminus*: 0.046, $SE = 0.006$; $p < .001$) ROHs is significantly greater in *S. catenatus* (Figure 5c). The latter increases indicate that more recent demographic events have had a substantially greater impact on *S. catenatus*.

Finally, *S. catenatus* populations show few differences, if any, in size class N_{ROH} profiles (Figure 5b; Figure S5a–c). Overall, 17%, 33% and 0% of N_{ROH} pairwise comparisons among populations are significant ($p < .05$) for short, medium and long ROH tracts, respectively (Figure 5b; Figure S5a–c). We obtained similar results using relative proportion metrics for each ROH-size class across *S. catenatus* populations (Figure 5d; Figure S5d–f). Broadly, this suggests similar magnitudes of population declines across different populations spanning the range of the species.

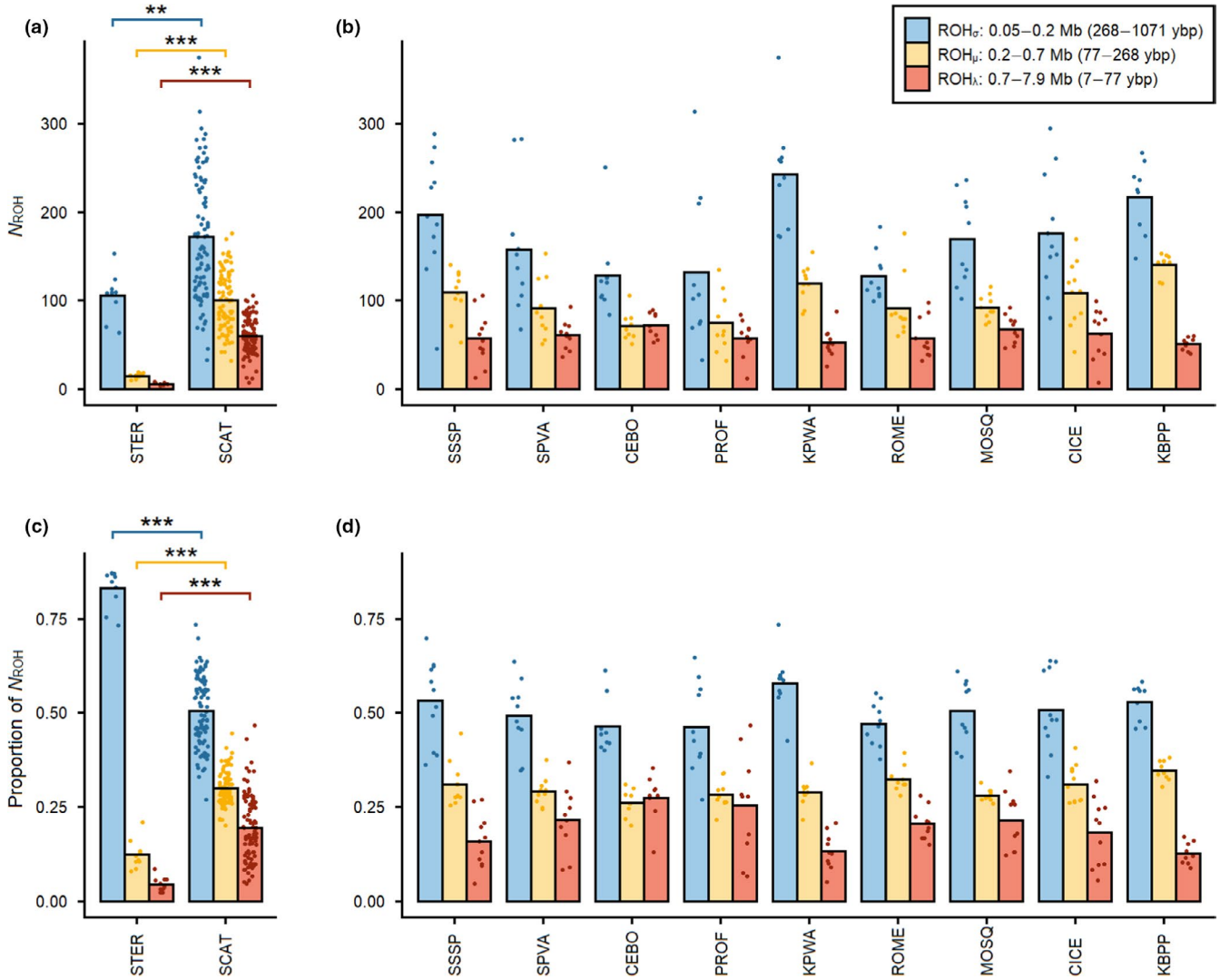

FIGURE 5 Mean N_{ROH} for short (ROH_{σ} = 0.05–0.2 Mb; blue bars), medium (ROH_{μ} = 0.2–0.7 Mb; yellow bars) and long (ROH_{λ} = 0.7–7.9 Mb; red bars) ROHs across individuals (circles) from the same (a) species (STER = *S. tergeminus*; SCAT = *S. catenatus*) and (b) *S. catenatus* population (SSSP, SPVA, CEBO, PROF, KPWA, ROME, MOSQ, CICE, KBPP). Mean relative N_{ROH} metrics for each (c) species and (d) *S. catenatus* population are also shown. Each ROH size class has an associated time of coalescence, which is included in the main legend as years before present (ybp). ROH-size class statistical tests are shown at the species level, where $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$ and ns = nonsignificant. Statistical tests between *S. catenatus* populations are shown in Figure S5 [Colour figure can be viewed at wileyonlinelibrary.com]

3.3 | Long-term demographic history

Simulations with PSMC (Figure 6) indicate that since the Middle Pleistocene (~500,000 ybp), *S. catenatus* has had a consistently smaller N_e than *S. tergeminus*. These results show that *S. tergeminus* reached a maximum N_e of ~190,000 individuals ~400,000 ybp, and had a stable N_e of ~65,000 individuals ~30,000–110,000 ybp, during the Late Pleistocene. After this time, *S. tergeminus* experienced declines in N_e, reaching a size of ~15,000 individuals ~10,000 ybp, in the Holocene. In contrast, *S. catenatus* reached a larger maximum N_e of ~230,000 individuals ~1,000,000 ybp, during the Early Pleistocene. This species then experienced consistent declines in N_e until reaching a stable size of ~20,000 individuals ~20,000–110,000 ybp, during the Late Pleistocene. Our results suggest that ~10,000 ybp, *S. catenatus* had an N_e of ~5000 individuals , which is three times smaller than that observed in *S. tergeminus* during this time.

3.4 | Mutation load

At the interspecific level (*S. catenatus* and *S. tergeminus* samples combined), we identified 27,419 sites containing derived nonsynonymous alleles from 3575 BUSCOs. Using PROVEAN, we characterized 91% (25,054) of these sites as benign, 5% (1402) as mildly deleterious, 3% (711) as moderately deleterious and 1% (252) as highly deleterious. Collectively, these represent species-level mutations, which can be used to estimate species-level mutation load.

Based on these sites, on average, the genomes of inbred *S. catenatus* individuals have lower absolute numbers and lower relative proportions of all classes of mutations compared to outbred *S. tergeminus* individuals ($p < .01$ in all cases; Figure 7a,b). Specifically, *S. catenatus* shows a 29% reduction in counts of benign alleles (*S. catenatus*: 4887, *SE* = 45; *S. tergeminus*: 6856, *SE* = 91), a 10% reduction in mildly deleterious alleles (*S. catenatus*: 181, *SE* = 2; *S. tergeminus*: 201, *SE* = 5), a 20% reduction in moderately deleterious alleles (*S. catenatus*: 65, *SE* = 1; *S. tergeminus*: 81, *SE* = 2) and, strikingly, a 62% reduction in highly deleterious alleles (*S. catenatus*: 24, *SE* < 1; *S. tergeminus*: 63, *SE* = 2; Figure 7a). Differences in the mean proportion of mutations of all classes are greatest for highly deleterious alleles (*S. catenatus*: 0.049, *SE* = 0.001; *S. tergeminus*: 0.145, *SE* = 0.003), intermediate for benign alleles (*S. catenatus*:

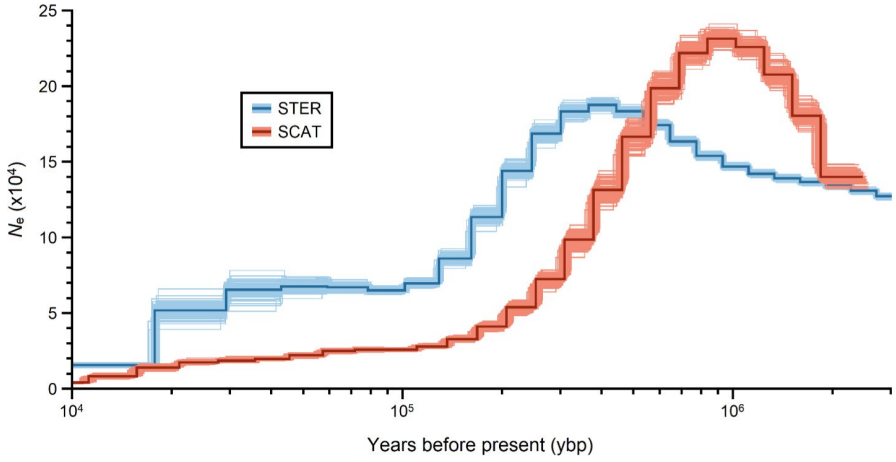

FIGURE 6 Historical changes in long-term effective population sizes (N_e) of *S. tergeminus* (dark blue lines) and *S. catenatus* (dark red lines) based on coalescent analyses. Light blue and light red lines reflect estimates across 100 bootstrap replicates for each species, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

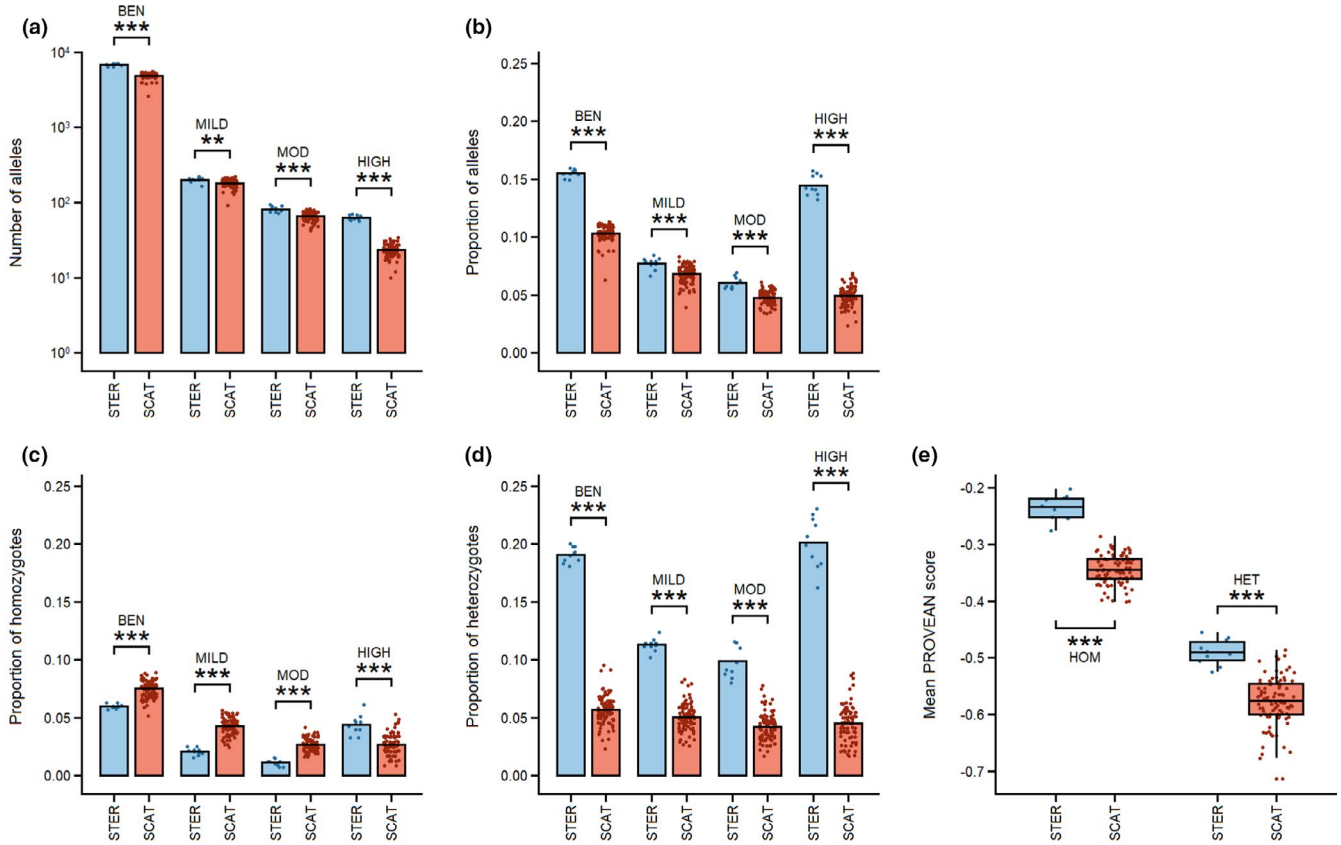

FIGURE 7 Distribution of benign (BEN), mildly deleterious (MILD), moderately deleterious (MOD) and highly deleterious (HIGH) alleles and genotypes for *S. tergeminus* (STER; blue circles) and *S. catenatus* (SCAT; red circles) individuals using different classes of species-level mutations. (a) Mean number of allele counts (bars) for each mutation class. Mean proportion (bars) of (b) alleles, (c) homozygotes and (d) heterozygotes for each mutation class. (e) Boxplots represent the distribution of mean PROVEAN scores across alleles present in homozygous (HOM) and heterozygous (HET) states. ***$p < .001$; **$p < .01$; *$p < .05$; ns = nonsignificant [Colour figure can be viewed at wileyonlinelibrary.com]

0.103, *SE* = 0.001; *S. tergeminus*: 0.155, *SE* = 0.001) and moderately deleterious alleles (*S. catenatus*: 0.048, *SE* = 0.001; *S. tergeminus*: 0.061, *SE* = 0.002) and smallest for mildly deleterious alleles (*S. catenatus*: 0.068, *SE* = 0.001; *S. tergeminus*: 0.077, *SE* = 0.002; Figure 7b).

Under the assumption that most mutations that arise in a population are recessive (Agrawal & Whitlock, 2011), the genotypic configuration of deleterious alleles can influence their realized negative impacts on individual fitness (Mathur & DeWoody, 2021). In this respect, *S. catenatus* has a significantly greater proportion of mildly and moderately deleterious homozygotes than *S. tergeminus* (Figure 7c), demonstrating greater levels of realized mutation load on a per-individual basis for two of the three classes of deleterious mutations examined in *S. catenatus*. The mean proportion of mildly and moderately deleterious homozygotes is 2.0-fold (*S. catenatus*: 0.043, *SE* = 0.001; *S. tergeminus*: 0.021, *SE* = 0.001) and 2.5-fold (*S. catenatus*: 0.027, *SE* = 0.001; *S. tergeminus*: 0.011, *SE* = 0.001) greater in *S. catenatus* than in *S. tergeminus*, respectively, while benign homozygotes are also found at a significantly greater mean frequency in *S. catenatus* (*S. catenatus*: 0.075, *SE* = 0.001; *S. tergeminus*: 0.060, *SE* = 0.001; Figure 7c). The exception are highly deleterious homozygotes, which are significantly less often found in *S. catenatus* (*S. catenatus*: 0.027, *SE* = 0.001; *S. tergeminus*: 0.044, *SE* = 0.003; Figure 7c). Differences are significant ($p < .001$) for all pairwise comparisons (Figure 7c). We interpret the latter comparison as reflecting purging of highly deleterious homozygotes in small *S. catenatus* populations, as has been found in small and repeatedly bottlenecked populations of Alpine ibex (Grossen et al., 2020; also see *Rxy* analyses below).

Although *S. tergeminus* shows higher deleterious allele counts than *S. catenatus* (Figure 7a), the fitness-related impacts of these alleles may be reduced because a greater proportion of all classes of deleterious mutations are found in heterozygous genotypes in *S. tergeminus* (Figure 7d). These differences are also significant for all four mutation classes ($p < .001$), including the benign class (Figure 7d).

Mean PROVEAN scores of derived alleles found in both homozygous and heterozygous genotypes are also significantly smaller (i.e., more deleterious) in *S. catenatus* than in *S. tergeminus* (Figure 7e). These results support the interpretation that high levels of drift present in small *S. catenatus* populations have allowed the persistence of mutations with greater potential negative impacts to fitness, regardless of genotype class.

In contrast, analyses that estimate population-level mutation load (Figure 8; Figures S6 and S7) show greater mean proportions of derived alleles in *S. catenatus* ($p < .001$ in all cases; Figure 8a), and a higher mean proportion of homozygous genotypes ($p < .001$ in all cases; Figure 8b). Patterns of population-level mutation load in heterozygous genotypes (Figure 8c) are less clear. As expected from the increased frequency of homozygotes, there is a significant ($p < .001$) decrease in the mean proportion of heterozygotes containing benign mutations in *S. catenatus* populations with respect to the representative *S. tergeminus* population examined, but no significant differences in mean proportions of heterozygotes for the three deleterious classes of mutations ($p > .05$; Figure 8c).

Comparisons of population-level mutation load across *S. catenatus* populations also show substantial and often significant differences in proportions of alleles and genotypes of all classes (Figures S6 and S7). Using the mean proportion of alleles from each mutation class as a measure of load (Figures S6a–d and S7a–d), we found >1.5-fold variations of benign (KPWA: 0.385, *SE* = 0.001; MOSQ: 0.584, *SE* = 0.001), mildly deleterious (SSSP: 0.240, *SE* = 0.005; MOSQ: 0.493, *SE* = 0.004), moderately deleterious (KPWA: 0.189, *SE* = 0.005; MOSQ: 0.477, *SE* = 0.010) and highly deleterious (KPWA: 0.232, *SE* = 0.008; MOSQ: 0.426, *SE* = 0.019) alleles across *S. catenatus* populations. These pairwise comparisons are significant ($p < .001$; Figures S6a–d and S7a–d) and are useful for identifying specific populations that have high (e.g., MOSQ) and low (e.g., KPWA and SSSP) mutation loads.

To assess possible drivers of population mutation load in *S. catenatus*, we correlated estimates of $N_{e.LT}$ (Table 1), short-term N_e ($N_{e.ST}$; Sovic et al., 2019; Table 1) and levels of inbreeding (mean F_{ROH}; Figure 3a) with mean proportions of alleles for each mutation class (Figure S6a–d; see Table S3 for correlations). Despite the small number

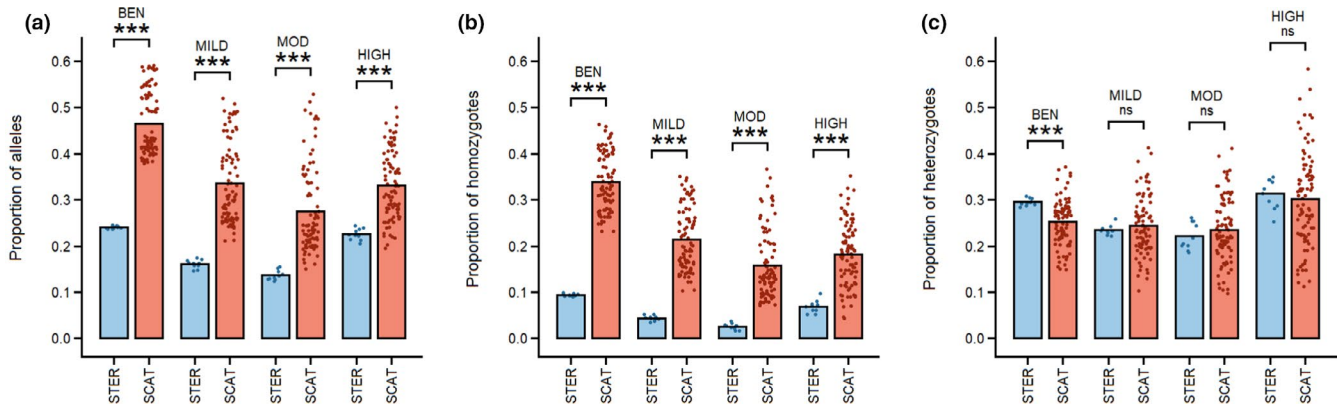

FIGURE 8 Distribution of benign (BEN), mildly deleterious (MILD), moderately deleterious (MOD) and highly deleterious (HIGH) alleles and genotypes for *S. tergeminus* (STER; blue circles) and *S. catenatus* (SCAT; red circles) individuals using different classes of population-level mutations. Mean proportion (bars) of (a) alleles, (b) homozygotes and (c) heterozygotes for each mutation class. ***$p < .001$; **$p < .01$; *$p < .05$; ns = nonsignificant [Colour figure can be viewed at wileyonlinelibrary.com]

of *S. catenatus* populations analysed, we found negative, and in some cases, significant ($p < .05$) or marginally nonsignificant ($p > .05$, but <.10) correlations between both measures of N_e and each mutation class, supporting the predicted impact of drift on population-level mutation load associated with population size (Table S3). In addition, we found positive correlations between mean F_{ROH} and mean proportion of alleles for each mutation class across *S. catenatus* populations, but none approached significance ($p > .05$ in all cases; Table S3).

Finally, *Rxy* analyses show that the relative frequencies of different mutation classes differ between inter- and intraspecific comparisons (Table 2). Both comparisons show similar relative frequencies of benign mutations between populations. However, for the interspecific comparison, the more deleterious mild and moderate mutations both show higher relative frequencies in the smaller *S. catenatus* population used to represent this species relative to *S. tergeminus* whereas the most harmful high class of mutations are less abundant in *S. catenatus*. In contrast, at the intraspecific level, comparisons between *S. catenatus* populations show that, with the exception of the benign mutations (see above), alleles that are classified into each of the three harmful mutation classes are found at a higher relative frequency in the smaller PROF population. Our interpretation is that these patterns represent differences in the impact of genetic purging on the frequency of the most deleterious mutations at the within- vs. between-species levels. Specifically, purging of highly deleterious mutations shapes differences in mutational load between these species, whereas there is no evidence that it operates to a significant extent between large and small populations of *S. catenatus*.

4 | DISCUSSION

4.1 | Demography

Demographic analyses indicate that increased inbreeding in *S. catenatus* is probably due to decreased population size relative to *S. tergenimus* over both long- and short-term timescales. Over evolutionary timescales (≥10,000 ybp), the PSMC analysis shows that for much of the last 500,000 years, *S. catenatus* has had an N_e three times smaller than that of *S. tergeminus*, even in the face of shifting distributions due to impacts of ice sheet expansion and contraction during the late Pleistocene (Pielou, 1992). Fundamental differences in the ecology of these species may account for the long-term differences in N_e. *S. catenatus* is a habitat specialist that occupies wetlands or mesic prairie habitats (Szymanski et al., 2016) which are patchily distributed, leading to greater population structure and smaller N_e sizes (Sovic et al., 2019). In contrast, *S. tergeminus* is found in more widespread continuous xeric grasslands (Stebbins, 1966), and this is reflected in the lack of genetic structure and large N_e sizes observed in nature (McCluskey & Bender, 2015).

Bottlenecks over recent timescales (thousands of years) have also increased inbreeding in *S. catenatus* relative to *S. tergeminus*. The oldest recent event that was detected occurred at timescales consistent with bottlenecks identified by Sovic et al. (2019) in *S. catenatus* using a different data set and analyses, and it had a proportionally greater impact on *S. tergeminus* based on the increased relative abundance of small ROHs in this species. The causes of this decline are unknown, but they seem most likely to be associated with large-scale environmental changes related to climate that have previously been hypothesized to impact the distribution of this and other species in this region of North America (Cook, 1992; Soltis et al., 2006).

In contrast, the two most recent bottleneck events occurred within the last 250 years, and the greater relative abundance of both medium and long ROHs in *S. catenatus* suggests that relative declines in population size have been greater in this species than in *S. tergeminus*. This time frame coincides with the colonization and subsequent landscape modification by European settlers in North America (Pielou, 1992; Schmidt, 1938), which has resulted in the habitat destruction and fragmentation that has played a key role in recent range-wide declines of *S. catenatus* (Szymanski et al., 2016). Overall, our analyses provide new historical evidence in support of the idea that anthropogenic impacts have had a consistently greater impact on *S. catenatus*.

4.2 | Mutation load

A key finding of our study is that analyses of mutations with different evolutionary histories are important for understanding the different ways in which mutation load is realized at species and population

TABLE 2 Relative mutation load estimated as *Rxy* for different mutation classes

Mutation class	Interspecific: SCAT-STER (*SE*; no. of scaffolds)	Intraspecific: PROF-KPWA (*SE*; no. of scaffolds)
Benign	1.07 (<0.01; 836)	1.02 (<0.01; 776)
Mildly deleterious	1.50 (<0.01; 174)	1.04 (<0.01; 174)
Moderately deleterious	1.34 (0.02; 72)	1.11 (0.01; 82)
Highly deleterious	0.53 (0.07; 24)	1.50 (0.06; 32)

Rxy values are presented for an interspecific comparison between *S. catenatus* (SCAT, represented by the KPWA population) and *S. tergeminus* (represented by the STER population) and for an intraspecific comparison between relatively small (PROF) and large (KPWA) *S. catenatus* populations. Values are presented such that *Rxy* >1 means that a given class allele has a higher frequency in the first population with respect to the second population of each pairwise comparison. Standard errors (*SE*) were obtained by jackknifing *Rxy* values associated with individual scaffolds (number also indicated) from each pairwise comparison.

levels. Specifically, we first estimated species-level mutation load using derived nonsynonymous alleles at sites that were polymorphic in either one or both species (species-level mutations). This criterion incorporates sites that reflect the almost two-fold difference in $\theta_{W.OUT}$ between species, which can substantially impact estimates of load (Figure 7). Based on these sites we found that, on average, the genomes of outbred *S. tergeminus* individuals had significantly more deleterious mutations compared to more inbred *S. catenatus* individuals when mutation load was estimated in terms of mean count or proportion of derived alleles (Figure 7a,b). Therefore, based on absolute numbers of deleterious alleles, the inbred threatened species (*S. catenatus*) has lower levels of mutation load than the more outbred common species (*S. tergeminus*).

However, the impact of deleterious mutations is dependent on the genotypic segregation patterns of deleterious alleles and the assumed dominance coefficients of mutations (Charlesworth, 2009). Empirical evidence suggests that many deleterious mutations are recessive and are mostly expressed if they segregate in homozygous genotypes (Agrawal & Whitlock, 2011). If this is the case, then the impact of mutation is lessened if they are mainly found in heterozygote genotypes. This is the case in outbred *S. tergeminus*: heterozygotes containing deleterious mutations are ~2–4-fold more common in *S. tergeminus* than in *S. catenatus* genomes (Figure 7d). This pattern argues that although *S. tergeminus* has a higher "potential" mutation load in the form of absolute numbers of deleterious mutations, the extent to which the negative effects of these mutations ("realized" load) occur is blunted by the fact that proportionately more occur in presumably unexpressed heterozygous genotypes and that they are also found at lower frequencies on a per-site basis (Mathur & DeWoody, 2021).

Finally, *Rxy* analyses provide another perspective on species mutation load by estimating load in terms of the frequency of specific classes of deleterious mutations relative to existing levels of neutral variation. Based on the populations compared, the results show that deleterious mutations with moderate negative impacts (mild and moderate categories) are at significantly higher frequencies in the more inbred *S. catenatus* population used in this comparison; but in striking contrast, the most damaging mutations (the high category) are significantly less abundant in *S. catenatus*. This pattern has been found in other studies comparing mutation load in inbred compared to outbred species (Grossen et al., 2020; Xue et al., 2015). This argues for a mechanism whereby species with reduced population sizes, and hence with greater susceptibilities to the effects of drift, accumulate moderately deleterious variants due to relaxed purifying selection, but experience genetic purging of highly deleterious mutations which are under strong purifying selection.

Overall, these results support the argument of van der Valk et al. (2021) that differences in mutation load between species need to take into account the impact of long-term demographic patterns on levels of deleterious mutations. Specifically, if threatened species have always persisted at low N_e sizes over evolutionary timescales, then sustained genetic purging of deleterious alleles can explain reduced mutation load in threatened species that currently exist in small populations. The long- and short-term demographic differences documented here support this explanation for the decreased mutation load shown in *S. catenatus* compared to *S. tergeminus*. The ROH analyses show that recent population bottlenecks have occurred in both species, with more recent events having a greater impact on *S. catenatus*. In addition, coalescent-based analyses demonstrate that *S. catenatus* has consistently had a smaller N_e over the past 500,000 years. These results argue that drift effects have been consistently stronger in the species with lower absolute levels of mutation load, which is consistent with the explanation that lower load is due to sustained purging of mutations over short and long timescales.

We also examined mutation load using derived mutations that were polymorphic within populations (population-level mutations) and thus reflect the impact of population-level processes. These mutations show strikingly different patterns from species-level mutations. Specifically, *S. catenatus* populations have, on average, higher proportions of these mutations on a per-genome basis across all impact classes, and all classes of mutations are found at higher levels in homozygote genotypes compared to *S. tergeminus*. This pattern is mirrored in derived mutation analyses based on *Rxy* values, which show higher frequencies of alleles for all mutation classes, including mutations from the class (high) predicted to be the most harmful. This indicates a lack of genetic purging of these mutations at the population level despite high levels of inbreeding. These patterns mirror those found in other threatened species with small populations (Robinson et al., 2016, 2019), and we interpret them as reflecting the negative impact of genetic drift on the efficiency with which natural selection removes deleterious mutations from small populations (Lohmueller, 2014). This interpretation is also supported by the fact that we find inverse and sometimes significant correlations between long- and short-term estimates of N_e and genetic load that signal an impact of genetic drift on variation in observed levels of load among *S. catenatus* populations.

It is unclear why we find a lack of evidence for genetic purging of highly deleterious mutations between large and small *S. catenatus* populations when there is a clear signal of this process taking place at the species level. A possible reason lies in differences in the timing and magnitude of inbreeding within and between species. Inbreeding is a key component of the purging mechanism because mating between relatives leads to increased representation of recessive mutations in homozygous genotypes ("realized load"; Mathur & DeWoody, 2021). The similarity of F_{ROH} values between *S. catenatus* populations (Figure 3) and lack of close relatives in most populations (Figure 4; also see Martin et al., 2021) suggests that levels of recent inbreeding are similar between populations and may only have occurred at a significant level relatively recently. Both factors may have reduced the impact of inbreeding on genetic purging of highly deleterious mutations at the population level compared to the species level in these snakes.

4.3 | Conservation implications

We see several implications from our results for managing genetic erosion in *S. catenatus* due to range-wide declines in population

numbers. First, at the species level, reduced genome-wide mutation load due to long-term purging suggests that future population declines in less impacted regions, such as the northern parts of the species range, will have fewer negative genetic impacts than comparable declines in threatened snakes with historically larger N_e sizes. *S. catenatus* may be more resistant to the process of mutational meltdown that is proposed to occur in small populations of threatened species (Frankham et al., 2017), leading to a focus on the greater importance of ecological factors to long-term population viability (Lande, 1988).

Second, consideration of the potential effects of long-term purging on mutational load has led to suggestions that conventional guidelines for conducting genetic rescue of small inbred populations of threatened species need to be reconsidered (Kyriazis et al., 2020; van der Valk et al., 2021; but see Ralls et al., 2020). Specifically, Kyriazis et al. (2020) argue that the typical approach for managing these populations is to maintain high genetic diversity through the transfer of individuals from large, genetically diverse populations, but that this carries a risk of introducing large numbers of deleterious mutations that can be exposed by inbreeding. Our empirical results suggest that, in fact, there is an inverse relationship between levels of mutational load and N_e among *S. catenatus* populations or that large potential donor populations have fewer, not more, potential deleterious mutations present. This observation, combined with the fact that donor individuals from larger populations will have greater numbers of potentially adaptive variants (Ochoa et al., 2020), broadly supports the idea that choosing donor individuals from populations with large N_e sizes is a sound strategy for genetic rescue in this species. This claim comes with two important caveats. It assumes a high degree of sharing of deleterious mutations between populations such that the interpopulation transfer of individuals will not introduce large numbers of new distinct negative mutations. Second, it also assumes that adaptive variation that evolves through local adaptation will have similar positive effects on individual fitness in populations receiving translocated individuals.

More broadly, our study provides an example of how whole genome sequences from threatened and endangered species can provide new approaches to assess patterns of functional variation that impact genetic erosion in these species (Leroy et al., 2018). When combined with the results of Ochoa et al. (2020), it provides a rare comprehensive assessment of two of the key aspects of genetic erosion in a single endangered species, namely the magnitude and evolutionary mechanisms shaping negative genetic load and adaptive genetic variation. As such it represents an important realization of the application of genomic tools and analyses for addressing genetic issues related to conservation of biodiversity at the species and population levels (Funk et al., 2019; Hohenlohe et al., 2021).

ACKNOWLEDGEMENTS

We thank all those individuals who have provided samples or assisted with collections across the range of *S. catenatus* over the past 25 years—this work would not have been possible without their help. Many individuals helped, but we especially thank Jeff Davis, Michael Dreslik, Brian Fedorko, Tony Frazier, Kim Frolich, Dan Harvey, Matt Kowalski, Greg Lipps, Chris Parent, Chris Phillips, Paul Pratt, Kent Prior, Kevin Shoemaker, Michelle Villenueve and Doug Wynn for their enthusiastic and generous assistance with finding snakes and/or providing samples. We thank Anna Brüniche-Olsen for introducing us to the idea of ROHs and encouraging us to think about analysing them in our snakes; Samarth Mathur, Scott Martin, Drew Schield and Todd Castoe for help with data and analyses; and Anna Brüniche-Olsen, J. Andrew DeWoody, Gideon Bradburd, Robert Fitak, Katerina Guschanski, Christine Grossen and Samarth Mathur for discussions and comments on the manuscript. Three anonymous reviewers offered comments that improved the manuscript. We also thank Kate Parsons and Carolyn Caldwell for their long-term support of our conservation genetics work on endangered snakes. Computational analyses were performed on the Ohio Biodiversity Conservation Partnership Computing Cluster using resources provided by the Ohio Supercomputer Center. This work was supported by the State Wildlife Grants Program, administered jointly by the U.S. Fish and Wildlife Service and the Ohio Division of Wildlife, with funds provided by the Ohio Biodiversity Conservation Partnership between The Ohio State University and the Ohio Division of Wildlife. H.L.G. was also supported by National Science Foundation (USA) Grant DEB 1638872 during the preparation of the manuscript.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

Designed research: A.O. and H.L.G. Performed research: A.O. and H.L.G. Analysed data: A.O. Wrote the paper: A.O. and H.L.G.

DATA AVAILABILITY STATEMENT

Raw sequence data used in this study are available in GenBank under BioProject PRJNA750087 (BioSample accessions SAMN20438760–859). The data that support the findings of this study (including code, input and output files, and statistical tests) are available in Dryad at: https://doi.org/10.5061/dryad.j9kd51cd5

ORCID

Alexander Ochoa https://orcid.org/0000-0002-3271-3684
H. Lisle Gibbs https://orcid.org/0000-0001-7461-3393

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Agrawal, A. F., & Whitlock, M. C. (2011). Inferences about the distribution of dominance drawn from yeast gene knockout data. *Genetics*, *187*(2), 553–566. https://doi.org/10.1534/genetics.110.124560

Allendorf, F. W. (2017). Genetics and the conservation of natural populations: Allozymes to genomes. *Molecular Ecology*, *26*(2), 420–430. https://doi.org/10.1111/mec.13948

Benazzo, A., Trucchi, E., Cahill, J. A., Maisano Delser, P., Mona, S., Fumagalli, M., Bunnefeld, L., Cornetti, L., Ghirotto, S., Girardi, M.,

Ometto, L., Panziera, A., Rota-Stabelli, O., Zanetti, E., Karamanlidis, A., Groff, C., Paule, L., Gentile, L., Vilà, C., … Bertorelle, G. (2017). Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proceedings of the National Academy of Sciences*, *114*(45), E9589–E9597. https://doi.org/10.1073/pnas.1707279114

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, *57*(1), 289–300. https://doi.org/10.2307/2346101

Bosse, M., Megens, H.-J., Madsen, O., Paudel, Y., Frantz, L. A. F., Schook, L. B., Crooijmans, R. P. M. A., & Groenen, M. A. M. (2012). Regions of homozygosity in the porcine genome: Consequence of demography and the recombination landscape. *PLoS Genetics*, *8*(11), e1003100. https://doi.org/10.1371/journal.pgen.1003100

Brüniche-Olsen, A., Kellner, K. F., Anderson, C. J., & DeWoody, J. A. (2018). Runs of homozygosity have utility in mammalian conservation and evolutionary studies. *Conservation Genetics*, *19*(6), 1295–1307. https://doi.org/10.1007/s10592-018-1099-y

Bylsma, R. R. (2020). *Population genetic and genomic analyses of western massasauga (*Sistrurus tergeminus *ssp.): Subspecies delimitation and conservation status* (Unpublished MS thesis). Purdue University.

Caughley, G. (1994). Directions in conservation biology. *Journal of Animal Ecology*, *63*, 215–244. https://doi.org/10.2307/5542

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews Genetics*, *19*(4), 220–234. https://doi.org/10.1038/nrg.2017.109

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*, 195–205. https://doi.org/10.1038/nrg2526

Chiucchi, J. E., & Gibbs, H. L. (2010). Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. *Molecular Ecology*, *19*(24), 5345–5358. https://doi.org/10.1111/j.1365-294X.2010.04860.x

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745–2747. https://doi.org/10.1093/bioinformatics/btv195

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, *7*(10), e46688. https://doi.org/10.1371/journal.pone.0046688

Cook, F. R. (1992). After an ice age: Zoogeography of the massasauga within a Canadian herpetological perspective. In B. Johnson, & V. Menzies (Eds.), *International symposium and workshop on the conservation of the eastern Massasauga Rattlesnake, Sistrurus catenatus catenatus* (pp. 19–25). Toronto Zoo.

Do, R., Balick, D., Li, H., Adzhubei, I., Sunyaev, S., & Reich, D. (2015). No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. *Nature Genetics*, *47*(2), 126–131. https://doi.org/10.1038/ng.3186

Frankham, R., Ballou, J. D., Ralls, K., Eldridge, M., Dudash, M. R., Fenster, C. B., Lacy, R. C., & Sunnucks, P. (2017). *Genetic management of fragmented animal and plant populations*. Oxford University Press.

Funk, W. C., Forester, B. R., Converse, S. J., Darst, C., & Morey, S. (2019). Improving conservation policy with genomics: A guide to integrating adaptive potential into U.S. Endangered Species Act decisions for conservation practitioners and geneticists. *Conservation Genetics*, *20*(1), 115–134. https://doi.org/10.1007/s10592-018-1096-1

García-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: Inbreeding, purging, mutation, and standard selection. *Genetics*, *190*(4), 1461–1476. https://doi.org/10.1534/genetics.111.135541

Gibbs, H. L., & Chiucchi, J. E. (2012). Inbreeding, body condition, and heterozygosity-fitness correlations in isolated populations of the endangered eastern massasauga rattlesnake (*Sistrurus c. catenatus*). *Conservation Genetics*, *13*(4), 1133–1143. https://doi.org/10.1007/s10592-012-0360-z

Glémin, S. (2003). How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution*, *57*(12), 2678–2687. https://doi.org/10.1111/j.0014-3820.2003.tb01512.x

Government of Canada. (2009). *Species at risk public registry*. Retrieved from https://www.sararegistry.gc.ca

Green, R. E., Braun, E. L., Armstrong, J., Earl, D., Nguyen, N., Hickey, G., Vandewege, M. W., St. John, J. A., Capella-Gutierrez, S., Castoe, T. A., Kern, C., Fujita, M. K., Opazo, J. C., Jurka, J., Kojima, K. K., Caballero, J., Hubley, R. M., Smit, A. F., Platt, R. N., … Ray, D. A. (2014). Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science*, *346*(6215), 12544–12549. https://doi.org/10.1126/science.1254449

Grossen, C., Guillaume, F., Keller, L. K., & Croll, D. (2020). Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nature Communications*, *11*(1), 1001. https://doi.org/10.1038/s41467-020-14803-1

Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding inbreeding depression, purging, and genetic rescue. *Trends in Ecology & Evolution*, *31*(12), 940–952. https://doi.org/10.1016/j.tree.2016.09.005

Hedrick, P. W., & Lacy, R. C. (2015). Measuring relatedness between inbred individuals. *Journal of Heredity*, *106*(1), 20–25. https://doi.org/10.1093/jhered/esu072

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62–82. https://doi.org/10.1111/mec.15720

International Chicken Genome Sequencing Consortium. (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, *432*(7018), 695–716. https://doi.org/10.1038/nature03154

Jain, C., Koren, S., Dilthey, A., Phillippy, A. M., & Aluru, S. (2018). A fast adaptive algorithm for computing whole-genome homology maps. *Bioinformatics*, *34*(17), i748–i756. https://doi.org/10.1093/bioinformatics/bty597

Jellen, B. C., Shepard, D. B., Dreslik, M. J., & Phillips, C. A. (2007). Male movement and body size affect mate acquisition in the eastern massasauga (*Sistrurus catenatus*). *Journal of Herpetology*, *41*(3), 451–457.

Kardos, M., Luikart, G., & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. *Heredity*, *115*(1), 63–72. https://doi.org/10.1038/hdy.2015.17

Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016). Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications*, *9*(10), 1205–1218. https://doi.org/10.1111/eva.12414

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230–241. https://doi.org/10.1016/S0169-5347(02)02489-8

Kirkpatrick, M., & Jarne, P. (2000). The effects of a bottleneck on inbreeding depression and the genetic load. *American Naturalist*, *155*(2), 154–167. https://doi.org/10.1086/303312

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Korneliussen, T. S., & Moltke, I. (2015). NgsRelate: A software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics*, *31*(24), 4009–4011. https://doi.org/10.1093/bioinformatics/btv509

Kyriazis, C. C., Wayne, R. K., & Lohmueller, K. E. (2020). Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. *Evolution Letters*, *5*(1), 33–47. https://doi.org/10.1002/evl3.209

Lande, R. (1988). Genetics and demography in biological conservation. *Science*, *241*(4872), 1455–1460. https://doi.org/10.1126/science.3420403

Leroy, G., Carroll, E. L., Bruford, M. W., DeWoody, J. A., Strand, A., Waits, L., & Wang, J. (2018). Next-generation metrics for monitoring genetic erosion within populations of conservation concern. *Evolutionary Applications*, *11*(7), 1066–1083. https://doi.org/10.1111/eva.12564

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, *475*(7357), 493–496. https://doi.org/10.1038/nature10231

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lohmueller, K. E. (2014). The impact of population demography and selection on the genetic architecture of complex traits. *PLoS Genetics*, *10*, e1004379. https://doi.org/10.1371/journal.pgen.1004379

Martin, S. A., Lipps, G. J. Jr., & Gibbs, H. L. (2021). Pedigree-based assessment of recent population connectivity in a threatened rattlesnake. *Molecular Ecology Resources*, *21*(6), 1820–1832. https://doi.org/10.1111/1755-0998.13383

Mathur, S., & DeWoody, J. A. (2021). Genetic load has potential in large populations but is realized in small inbred populations. *Evolutionary Applications*, *14*(6), 1540–1557. https://doi.org/10.1111/eva.13216

McCluskey, E. M., & Bender, D. (2015). Genetic structure of western massasauga rattlesnakes (*Sistrurus catenatus tergeminus*). *Journal of Herpetology*, *49*(3), 343–348. https://doi.org/10.1670/14-016

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 254–260. https://doi.org/10.1101/gr.107524.110.20

Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5269–5273. https://doi.org/10.1073/pnas.76.10.5269

Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012). SNP calling, genotype calling, and sample allele frequency estimation from New-Generation Sequencing data. *PLoS One*, *7*(7), e37558. https://doi.org/10.1371/journal.pone.0037558

Ochoa, A., Broe, M., Moriarty Lemmon, E., Lemmon, A. R., Rokyta, D. R., & Gibbs, H. L. (2020). Drift, selection and adaptive variation in small populations of a threatened rattlesnake. *Molecular Ecology*, *29*(14), 2612–2625. https://doi.org/10.1111/mec.15517

Pasquesi, G. I. M., Adams, R. H., Card, D. C., Schield, D. R., Corbin, A. B., Perry, B. W., Reyes-Velasco, J., Ruggiero, R. P., Vandewege, M. W., Shortt, J. A., & Castoe, T. A. (2018). Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. *Nature Communications*, *9*(1), 2774. https://doi.org/10.1038/s41467-018-05279-1

Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics*, *91*(2), 275–292. https://doi.org/10.1016/j.ajhg.2012.06.014

Perrier, C., Ferchaud, A. L., Sirois, P., Thibault, I., & Bernatchez, L. (2017). Do genetic drift and accumulation of deleterious mutations preclude adaptation? Empirical investigation using RADseq in a northern lacustrine fish. *Molecular Ecology*, *26*(22), 6317–6335. https://doi.org/10.1111/mec.14361

Pielou, E. C. (1992). *After the ice age: The return of life to glaciated North America*. University of Chicago Press.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ralls, K., Sunnucks, P., Lacy, R. C., & Frankham, R. (2020). Genetic rescue: A critique of the evidence supports maximizing genetic diversity rather than minimizing the introduction of putatively harmful genetic variation. *Biological Conservation*, *251*, 108784. https://doi.org/10.1016/j.biocon.2020.108784

Renaud, G., Hanghøj, K., Korneliussen, T. S., Willerslev, E., & Orlando, L. (2019). Joint estimates of heterozygosity and runs of homozygosity for modern and ancient samples. *Genetics*, *212*(3), 587–614. https://doi.org/10.1534/genetics.119.302057

Renaut, S., & Rieseberg, L. H. (2015). The accumulation of deleterious mutations as a consequence of domestication and improvement in sunflowers and other Compositae crops. *Molecular Biology and Evolution*, *32*(9), 2273–2283. https://doi.org/10.1093/molbev/msv106

Robinson, J. A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B. Y., vonHoldt, B. M., Marsden, C. D., Lohmueller, K. E., & Wayne, R. K. (2016). Genomic flatlining in the endangered island fox. *Current Biology*, *26*(9), 1183–1189. https://doi.org/10.1016/j.cub.2016.02.062

Robinson, J. A., Räikkönen, J., Vucetich, L. M., Vucetich, J. A., Peterson, R. O., Lohmueller, K. E., & Wayne, R. K. (2019). Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Science Advances*, *5*(5), eaau0757. https://doi.org/10.1126/sciadv.aau

Ryberg, W. A., Harvey, J. A., Blick, A., Hibbitts, T. J., & Voelker, G. (2015). Genetic structure is inconsistent with subspecies designations in the western massasauga *Sistrurus tergeminus*. *Journal of Fish and Wildlife Management*, *6*(2), 350–359. https://doi.org/10.3996/122014-JFWM-093

Saremi, N. F., Supple, M. A., Byrne, A., Cahill, J. A., Coutinho, L. L., Dalén, L., Figueiró, H. V., Johnson, W. E., Milne, H. J., O'Brien, S. J., O'Connell, B., Onorato, D. P., Riley, S. P. D., Sikich, J. A., Stahler, D. R., Villela, P. M. S., Vollmers, C., Wayne, R. K., Eizirik, E., … Shapiro, B. (2019). Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, *10*(1), 4769. https://doi.org/10.1038/s41467-019-12741-1

Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake recombination landscapes are concentrated in functional regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294. https://doi.org/10.1093/molbev/msaa003

Schmidt, K. P. (1938). Herpetological evidence for the post-glacial extension of the steppe in North America. *Ecology*, *19*, 396–407.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smit, A. F., Hubley, R., & Green, P. (1996– 2010). RepeatMasker Open-3.0. http://www.repeatmasker.org

Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, *15*(14), 4261–4293. https://doi.org/10.1111/j.1365-294X.2006.03061.x

Sovic, M. G., Fries, A. C., & Gibbs, H. L. (2016). Origin of a cryptic lineage in a threatened reptile through isolation and historical hybridization. *Heredity*, *117*(5), 358–366. https://doi.org/10.1038/hdy.2016.56

Sovic, M., Fries, A., Martin, S. A., & Gibbs, H. L. (2019). Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, *Sistrurus catenatus*. *Evolutionary Applications*, *12*(4), 664–678. https://doi.org/10.1111/eva.12731

Stebbins, R. C. (1966). *A field guide to western reptiles and amphibians*. Houghton Mifflin Co.

Stedman, A. L., Jaeger, C. P., Hileman, E. T., Jellen, B. C., Phillips, C. A., Swanson, B. J., & King, R. B. (2016). Multiple paternity in three wild populations of eastern massasauga (*Sistrurus catenatus*). *Herpetological Conservation and Biology*, *11*(1), 160–167.

Szymanski, J., Pollack, C., Ragan, L., Redmer, M., Clemency, L., Voorhies, K., & Jaka, J. (2016). *Species status assessment for the eastern massasauga rattlesnake (*Sistrurus catenatus*)*. U.S. Fish and Wildlife Service, Region 3.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595. https://doi.org/10.1093/genetics/123.3.585

U.S. Fish and Wildlife Service. (2016). Endangered and threatened wildlife and plants; threatened species status for the eastern massasauga rattlesnake. *Federal Register*, *81*(190), 67193–67214.

van der Valk, T., de Manuel, M., Marques-Bonet, T., & Guschanski, K. (2021). Estimates of genetic load suggest frequent purging of deleterious alleles in small populations. *bioRxiv*, 696831. https://doi.org/10.1101/696831

van Oosterhout, C. (2020). Mutation load is the spectre of species conservation. *Nature Ecology & Evolution*, *4*(8), 1004–1006. https://doi.org/10.1038/s41559-020-1204-8

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*(2), 256–276. https://doi.org/10.1016/0040-5809(75)90020-9

Xue, Y., Prado-Martinez, J., Sudmant, P. H., Narasimhan, V., Ayub, Q., Szpak, M., Frandsen, P., Chen, Y., Yngvadottir, B., Cooper, D. N., de Manuel, M., Hernandez-Rodriguez, J., Lobon, I., Siegismund, H. R., Pagani, L., Quail, M. A., Hvilsom, C., Mudakikwa, A., Eichler, E. E., … Scally, A. (2015). Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science*, *348*(6231), 242–245. https://doi.org/10.1126/science.aaa3952

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ochoa, A., & Gibbs, H. L. (2021). Genomic signatures of inbreeding and mutation load in a threatened rattlesnake. *Molecular Ecology*, 30, 5454–5469. https://doi.org/10.1111/mec.16147