PNAS RESEARCH ARTICLE | EVOLUTION

Functional genomic diversity is correlated with neutral genomic diversity in populations of an endangered rattlesnake

Samarth Mathur[a,b], Andrew J. Mason[a,b], Gideon S. Bradburd[c,d], and H. Lisle Gibbs[a,b,1]

Edited by Nils Stenseth, Universitetet i Oslo, Oslo, Norway; received February 22, 2023; accepted September 19, 2023

Theory predicts that genetic erosion in small, isolated populations of endangered species can be assessed using estimates of neutral genetic variation, yet this widely used approach has recently been questioned in the genomics era. Here, we leverage a chromosome-level genome assembly of an endangered rattlesnake (*Sistrurus catenatus*) combined with whole genome resequencing data (N = 110 individuals) to evaluate the relationship between levels of genome-wide neutral and functional diversity over historical and future timescales. As predicted, we found positive correlations between genome-wide estimates of neutral genetic diversity (π) and inferred levels of adaptive variation and an estimate of inbreeding mutation load, and a negative relationship between neutral diversity and an estimate of drift mutation load. However, these correlations were half as strong for projected future levels of neutral diversity based on contemporary effective population sizes. Broadly, our results confirm that estimates of neutral genetic diversity provide an accurate measure of genetic erosion in populations of a threatened vertebrate. They also provide nuance to the neutral-functional diversity controversy by suggesting that while these correlations exist, anthropogenetic impacts may have weakened these associations in the recent past and into the future.

neutral and functional genomic diversity | conservation genomics | *Sistrurus catenatus* | mutation load | adaptive variation

Significance

A cornerstone concept in the field of conservation genetics is that the amount of genetic erosion in small populations of endangered species can be assessed using estimates of neutral genetic diversity, but this idea has been questioned. We used genomic data to show that neutral diversity can provide an accurate measure of both components of functional genetic diversity—mutation load and adaptive variation—in populations of an endangered rattlesnake but that this correlation is weaker for projections of future levels of neutral diversity versus historical estimates. Our findings suggest a more nuanced perspective to the neutral-functional diversity controversy by suggesting that although these correlations are present, anthropogenetic impacts may have weakened these associations in present-day and future populations.

Threatened and endangered species that live in small, isolated populations can suffer elevated risk of extinction due to genetic erosion, which can lead to reduced individual fitness and ultimately population extirpation (1–3). Genetic erosion occurs when population numbers shrink, which results in a loss of adaptive variation through increased genetic drift, increased expression of genetic load through inbreeding, and the reduced efficiency of natural selection to remove harmful mutations (4). These impacts compromise the present and future viability of small populations through the negative effect of genetic load on individual fitness and the reduced ability of populations to adapt to environmental changes, including human-mediated stressors. Therefore, assessing the degree to which small populations suffer from genetic erosion is an important component of their management (5–7). One increasingly common approach which is facilitated by the increasing availability of genomic information across diverse taxa is to use molecular genetic proxies of positive and negative functional variation (8). This approach is challenging because it requires identifying specific genetic variants that are potentially linked to fitness in natural populations that in turn have demographic impacts—an often sought but rarely achieved goal in the fields of conservation genetics and evolutionary biology (9, 10). Instead, conservation geneticists have long used neutral genetic diversity or related demographic estimates like effective population size as surrogate measures of the amount of genetic erosion in small populations, under the assumption that neutral genetic diversity is correlated with functional genetic diversity and therefore predicts population viability (11–14).

This long-standing paradigm has been recently challenged by empirical results and theoretical considerations that suggest that there is no simple relationship between neutral and functional genetic diversity (15–18). For example, empirical studies find only weak links between surrogate measures of neutral and functional genetic diversity (19, 20), and there are small populations of endangered species with exceptionally low levels of overall genetic diversity that persist through time (21–23). These empirical findings may result from the fact that repeated historical bottlenecks can lead to purging of deleterious mutations from small populations, leading to reduced levels of genetic load (24–26). Recent theoretical work suggests that adaptive potential of small populations may be only weakly correlated with levels of neutral genetic diversity due to interactions between environmental variation and shifts in selection intensity acting on preexisting mutations (18). Finally, the expression of genetic load may not necessarily strongly impact mean absolute fitness due to density dependent fitness effects and the potentially strong dependence of mean absolute fitness on environmental conditions (27–30).

Author contributions: S.M., A.J.M., G.S.B., and H.L.G. designed research; S.M. and A.J.M. performed research; G.S.B. and H.L.G. contributed new reagents/analytic tools; S.M. and A.J.M. analyzed data; and S.M., A.J.M., G.S.B., and H.L.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: gibbs.128@osu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2303043120/-/DCSupplemental.

Published October 16, 2023.

Together, these empirical and theoretical arguments undermine the use of neutral genetic diversity or related measures as predictors of the functional diversity of a population, and hence of the population viability of threatened species that exist in small populations in the wild. However, these arguments have been countered by a multitude of studies emphasizing the robustness of the link between neutral genetic diversity and surrogate measures of fitness (14) and recent work demonstrating links between components of functional and neutral diversity at the interspecific level over historical timescales (8). Consequently, the nature of the relationship between neutral and functional genetic diversity in small populations of threatened species remains an important unresolved question in conservation genetics.

Specific predictions about the links between nucleotide diversity (π), genetic load, and adaptive variation are available from theory and population simulations (e.g., see Fig. 1 in ref. 30), although the strength of these correlations can weaken either due to true biological effects, as occurs in the wake of a population bottleneck, or as an artifact of how simulation models are parameterized (30). It is important to recognize that commonly used measures of neutral genetic diversity such as nucleotide diversity (π) partially reflect the strength of drift experienced by a population averaged over relatively long timescales (~$4N_e$ generations; where N_e is the effective population size) (31, 32), while information useful for conservation decisions often comes from shorter time scales (e.g., 10's of generations) (31, 33). Approaches that take advantage of genome-scale data offer the chance to accurately estimate contemporary N_e, which enables us to estimate expected levels of neutral genetic diversity over present-day and future timescales that are more relevant to conservation (34).

A valuable contribution would be to directly assess links between functional and neutral genetic diversity in populations of endangered species using more comprehensive genomic datasets and methods for identifying functional genetic diversity at the genome-wide level. The recent explosion in the availability of high-quality reference genomes and population resequencing data for endangered species, along with new methods for analyzing genomic data, raises the possibility of testing these predictions for individual species (36). A major challenge in testing the relationship between neutral and functional genetic diversity is obtaining plausible estimates of different components of functional genetic diversity (37–39). Progress has been made in estimating proxies of genetic load in natural populations of threatened species (for review, see ref. 40), and methods exist for obtaining estimates of adaptive variation at the molecular level in nonmodel species, although neither of these approaches directly measure the impacts of specific variants on fitness (10).

To meet this challenge, we adopt a "reverse ecology" approach that uses genome-wide selection scans based on gene-specific ratios of missense and silent polymorphisms and substitutions (35, 41) to identify genes that potentially contribute to variation in important selected phenotypes and are the targets of strong positive or negative selection (42). We suggest that variation in the form of nonsynonymous polymorphisms in coding sequences of genes under strong purifying selection along with impactful mutations like Loss of Function (LoF) mutations can be used to estimate genetic load and that nonsynonymous substitutions in genes identified by selection tests to be under strong positive selection represent potential adaptive diversity. Generating such data would enable empirical tests of theoretical predictions about how components of functional genetic diversity vary with historical and future estimates of neutral genetic diversity in small populations.

Here, we apply these approaches to assess the relationship between neutral and functional genetic diversity in small, isolated populations of the Eastern Massasauga rattlesnake (*Sistrurus catenatus*). Historically, this snake was found in wetlands and surrounding upland habitat in midwestern and eastern North America.

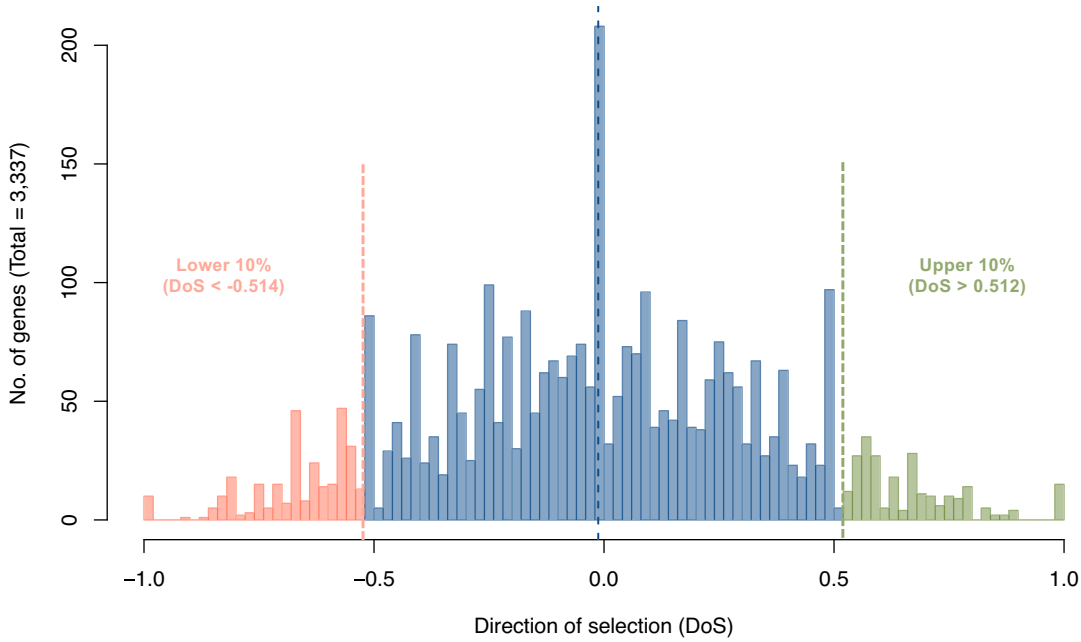

Fig. 1. Distribution of genome-wide gene-specific DoS values for 3,337 protein-coding genes identified in the annotated genome of *S. catenatus*. Values are based on comparisons of polymorphism between *S. catenatus* and *S. tergeminus* (*Materials and Methods*). A DoS value > 0 is evidence for a gene under positive (Darwinian) selection in *S. catenatus* while a DoS value of < 0 suggests a gene under negative (purifying) selection. The dotted line indicates the mean value (DoS = −0.0125). Cutoffs for the DoS values for genes with the highest (upper 10%; green) and lowest (lower 10%; red) DoS values are indicated. We used these categories to identify genes in which nonsynonymous substitutions represent potential adaptive diversity (upper 10%) or mutational load (lower 10%). A list of the genes that fall into each of these categories is given in Dataset S2. DoS values were calculated using the correction to the Neutrality Index suggested by ref. 35 (see *Materials and Methods* for details).

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 (43) and as a Species at Risk in Canada (44). This species exhibits little phylogeographic structure across its range—hence no evolutionary significant conservation units (45)—and high levels of population genetic structure (45–48), so that the relevant management units within this species are likely individual populations. Previous work has shown that drift impacts levels of adaptive variation in the form of nonsynonymous variation in snake venom gene coding sequences (49), suggesting a possible link between levels of neutral and putatively adaptive variation at the molecular level. However, there is also strong evidence that these populations have experienced recent population bottlenecks, raising questions about whether they are in genetic equilibrium over recent timescales (26, 48). Analysis of population genomic data offers a way to explore these questions both for this snake and, more generally, for species of conservation concern found in small populations in the wild.

In this study, we combine a recently sequenced and annotated genome for *S. catenatus* with population genome resequencing data to analyze the relationship between genome-wide estimates of functional genetic diversity and measures of drift intensity at both historic timescales and the more recent timescales that may be more relevant to conservation decisions. Our study represents an empirical assessment of the link between neutral and functional genetic diversity in a threatened species and provides a template for genome-level assessments of this fundamental relationship that is a cornerstone of the field of conservation genetics (18, 30).

Results

Genome Assembly and Annotation. Our final assembly for *S. catenatus* consisted of a 1.52-GB genome recovered in 2,004 scaffolds with a scaffold N50 of 195.71 MB and a contig N50 of 0.81 MB. Overall, this assembly was comparable in quality to other recently published viper genomes (*SI Appendix*, Table S1). BUSCO completeness was calculated as 91.1% complete, with 4.4% fragmented and 4.5% missing BUSCOs. In total, we recovered 17,061 annotated genes. Genome-wide variation in gene density, GC content, and repeat content are shown in *SI Appendix*, Fig. S1. Average gene density in 100-KB windows was 1.02 (1.14 among chromosome-length scaffolds). Average GC content across windows was 39.8% (range 25.9 to 72.4%), and average repeat content was 44.5%.

Identification of Protein-coding Genes under Positive and Negative Selection. We used divergence and polymorphism data from the largest population of *S. catenatus* in the study (KLDR; *SI Appendix*, Table S2) and a population of its sister species (*Sistrurus tergeminus*) to calculate Direction of Selection [DoS; (35)] values for protein-coding genes annotated in the *S. catenatus* genome. After excluding genes with low numbers of polymorphic and divergent sites (*Materials and Methods*), we estimated DoS for 3,337 genes (Fig. 1). The distribution of values was approximately normal and centered on a value slightly less than 0 (DoS = −0.0125 ± 0.365; mean ± SD). This suggests most coding sequences are evolving under neutral or nearly neutral processes with slightly more genes evolving under purifying than positive selection. Next, we classified genes as being under strong positive (Darwinian) or negative (purifying) selection by identifying genes with DoS values that lie within the upper 10% (DoS > 0.512) or lower 10% (DoS < −0.514) of the overall DoS distribution, respectively (N = 334; Fig.1 and Dataset S1). A gene ontology term analyses of the genes under positive selection identified cytoplasmic enzyme genes that regulate kinase activity and various cellular processes via metal ion binding (*SI Appendix*, Table S3 and *Supplementary Methods*), but also genes like *tlr2* and *tgfbr2* with known roles in the innate and adaptive immune systems respectively (Dataset S2). Most of the genes under purifying selection had function in olfactory reception and neural development (*SI Appendix*, Table S3 and Dataset S2).

Estimation of Functional and Neutral Diversity within and among *S. catenatus* Populations. We used whole genome data from 117 individuals from 12 isolated populations (mean no. individuals per population ± SD = 9.69 ± 1.49; *SI Appendix*, Fig. S2 and Table S2 and Dataset S3) to estimate different components of functional and neutral genetic diversity (*SI Appendix*, Tables S4 and S5). To generate estimates of different components of genetic load, we classified nonsynonymous substitutions in the lower 10% loci of the DoS distribution as deleterious mutations. We also included LoF mutations as deleterious because of their likely negative impact on fitness (50). Measures of genetic load require estimates of the fitness impacts of specific mutations, which were unavailable to us. As such, our mutation-based estimates are more appropriately described as estimates of mutation load defined in terms of the frequency and number of putatively deleterious mutations; we assume that these estimates are associated with true values of genetic load. To quantify adaptive variation, we classified nonsynonymous substitutions in genes in the upper 10% of the DoS distribution as potentially adaptive mutations. Next, for each population, we calculate $\pi_{adaptive}$, which is the nucleotide diversity at the nonsynonymous SNPs (single nucleotide polymorphisms) within the protein-coding regions of genes in the upper 10% of the DoS distribution. We used $\pi_{adaptive}$, as an estimate of the amount of additive diversity at a trait where diversity impacts the response of the population to present or future selection (30). We then used putatively deleterious variants to estimate mutation load and putatively adaptive variants to estimate adaptive diversity (see *SI Appendix*, Table S4 for details about each metric); we used these estimates together to quantify functional genetic diversity in *S. catenatus* populations and individuals (*SI Appendix*, Table S5).

Mutation load at the population level was estimated in two ways: first, the total number of deleterious mutations segregating within a population (N_{del}), and, second, the number of deleterious mutations that are at high frequency within the population [MAF (minor allele frequency) > 0.9; L_{drift}]. KLDR had the highest number of deleterious mutations segregating within the population (N_{del} = 2001 ± 42) but only ~1% of those deleterious mutations are nearly fixed (L_{drift} = 21.23 ± 7.7) whereas, JENN had the lowest number of segregating deleterious mutations (N_{del} = 1421) with ~11% of those are nearly fixed (L_{drift} = 162). At the individual level, mutation load was partitioned into inbreeding load ($L_{inbreed}$) and realized load ($L_{realized}$) based on the total number of heterozygous and homozygous genotypes for putatively deleterious variants, respectively, that are present within each individual genome. Note that individual estimates of load do not account for population allele frequencies and dominance coefficients of deleterious variant, both of which impact the actual inbreeding load in a population (see equation 3 in ref. 40). Individuals from the KLDR population had the highest number of deleterious loci in the heterozygous state ($L_{inbreed}$ = 797 ± 226), while individuals from the CEBO population had the highest number of deleterious alleles in the homozygote state ($L_{realized}$ = 388 ± 77). Adaptive diversity ($\pi_{adaptive}$) was based on functional diversity in putatively adaptive genes, and the KLDR population had the highest ($\pi_{adaptive}$ = 4.13 ± 1.69 × 10^{-4}) and the JENN population had the lowest ($\pi_{adaptive}$ = 2.15 ± 1.27 × 10^{-4}) adaptive diversity at a level which is almost half than that observed in KLDR.

We estimated levels of neutral genetic diversity (or projections thereof) for each population with measures that reflect either historical processes acting over thousands of generations or predict future levels of diversity once populations reach mutation-drift equilibrium. Historical estimates of neutral diversity were measured as nucleotide diversity within gene-desert regions ($\pi_{geneDes}$), i.e., genomic regions that are at least 500 kb upstream or downstream from an annotated gene which minimized physical genetic linkage as a cause of correlations between neural and functional variants. KLDR had the highest historical neutral genetic diversity ($\pi_{geneDes} = 2.72 \pm 2.51 \times 10^{-3}$) and JENN had the lowest historical neutral diversity ($\pi_{geneDes} = 1.68 \pm 2.1 \times 10^{-3}$). For a projected estimate of neutral genetic diversity in populations at future equilibrium, we estimated effective population size using the degree of population-level linkage disequilibrium (Ne_{LD}) between physically unlinked neutral SNPs in the gene deserts regions. KLDR had the highest contemporary effective population size ($Ne_{LD} = 65.84 \pm 1.28$), whereas CEBO had the lowest contemporary size ($Ne_{LD} = 4.38 \pm 0.09$) (*SI Appendix*, Table S5). The two estimates of neutral genetic diversity are weakly but significantly positively correlated with each other ($R^2 = 0.34$; $P = 0.03$).

Correlations between Estimates of Functional Diversity and Historical and Future Neutral Diversity. Our main result was that estimates of functional genetic diversity were, in general, significantly correlated in predicted ways with estimates of neutral genetic diversity, but that these correlations were stronger for historical estimates of neutral diversity relative to future projections of levels of neutral variation based on contemporary effective size.

For historical population-level estimates of neutral diversity, there were significant positive relationships between neutral nucleotide diversity ($\pi_{geneDes}$) and the total number of deleterious mutations segregating within a population (N_{del}) (Fig. 2*A*) and adaptive nucleotide diversity ($\pi_{adaptive}$) (Fig. 2*E*), and a significant negative relationship between $\pi_{geneDes}$ and the number of deleterious mutations with a frequency of > 0.9 in a population (drift mutation load; L_{drift}). All R^2 values for these relationships were ≥ 0.68 and highly significant ($P < 0.001$). Individual-based measures of inbreeding ($L_{inbreed}$) (Fig. 2*C*) and drift ($L_{realized}$) (Fig. 2*D*) mutation load also show the expected relationships with the exception that the relationship between neutral nucleotide diversity and individual drift mutation load, while negative as predicted, is nonsignificant ($P = 0.11$). Finally, consistent with the individual-based results, analyses using mixed-effect linear models with population as a fixed effect showed a significant positive relationship between neutral diversity and individual inbreeding mutation load ($L_{inbreed}$) ($R^2 = 0.11$; $P = 0.004$) and a nonsignificant negative relationship between individual drift mutation load ($R^2 = 0.02$; $P = 0.19$; *SI Appendix*, Table S6).

Estimates of neutral diversity based on contemporary effective population size (Ne_{LD}) showed similar associations with estimates of functional diversity, but the associations were approximately half as strong as for historical estimates based on R^2 values (Fig. 3). Specifically, there were moderate but significant positive relationships between estimates of Ne_{LD} and population-level inbreeding mutation load (N_{del}) ($R^2 = 0.35$; $P = 0.02$) and adaptive diversity ($\pi_{adaptive}$) ($R^2 = 0.38$; $P = 0.02$) and a significant negative relationship with population-level estimates of drift mutation load (L_{drift}) ($R^2 = 0.27$; $P = 0.048$). Likewise, individual inbreeding mutation load ($L_{inbreed}$) and individual realized load ($L_{realized}$) showed predicted relationships with Ne_{LD} ($R^2 \geq 0.29$; $P \leq 0.042$). Mixed-effect linear models with population included as a fixed effect showed marginally significant relationships between Ne_{LD} and $L_{inbreed}$ ($R^2 = 0.06$; $P = 0.05$), and Ne_{LD} and $L_{realized}$ ($R^2 = 0.05$; $P = 0.04$).

Finally, to verify observed differences in R^2 values between analyses using the two neutral genetic diversity metrics, we compared

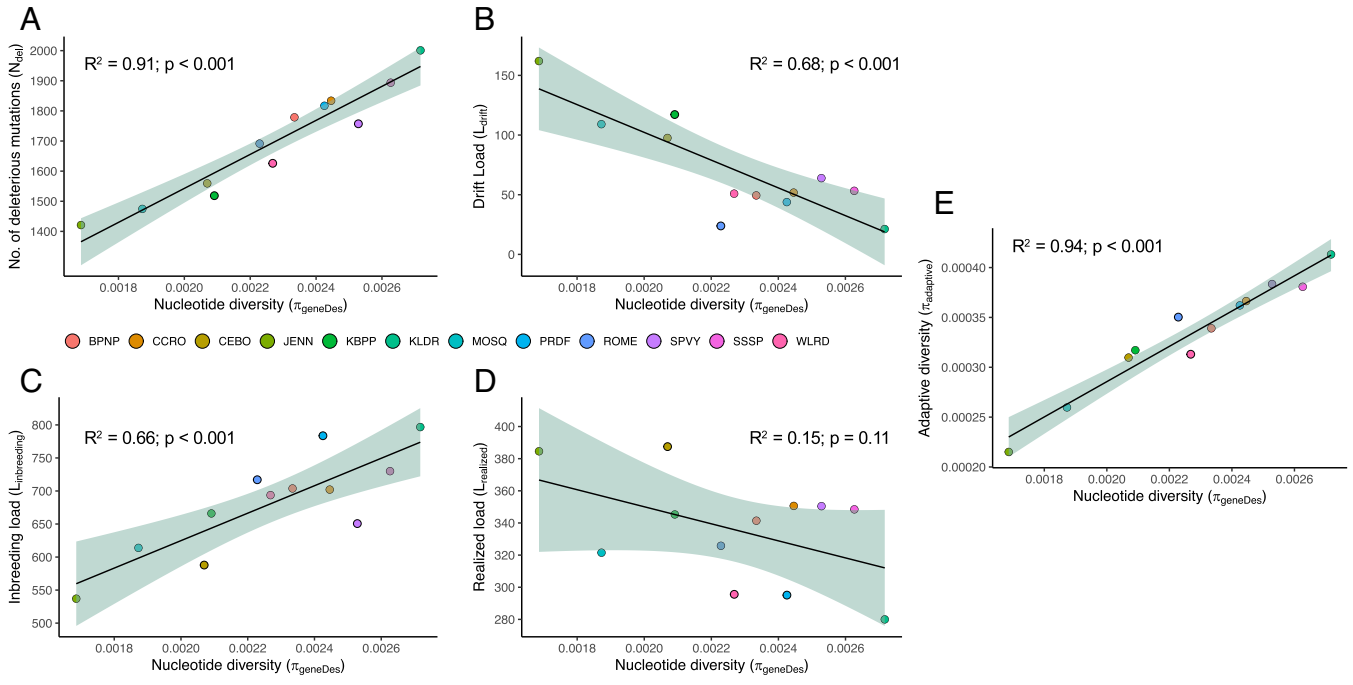

Fig. 2. Relationships between population estimates of functional and historical neutral genetic diversity. For each population, functional diversity was measured as (*A*–*D*) mutation load at a (*A* and *B*) population level or at an (*C* and *D*) individual level and as (*E*) adaptive diversity ($\pi_{adaptive}$) at a population level. Historical neutral genetic diversity was measured as mean nucleotide diversity in nonoverlapping 1-kb windows within gene deserts ($\pi_{geneDes}$) for each population. See *Materials and Methods* and *SI Appendix*, Table S4 for how each variable was calculated from *S. catenatus* population genomic data. The color-coded population origin for each data point is given in the legend. The black line represents the linear regression model, with the shaded region representing 95% CI. R^2 indicates the explained variance, and the *P*-value represents the statistical significance of the slope of the linear model.

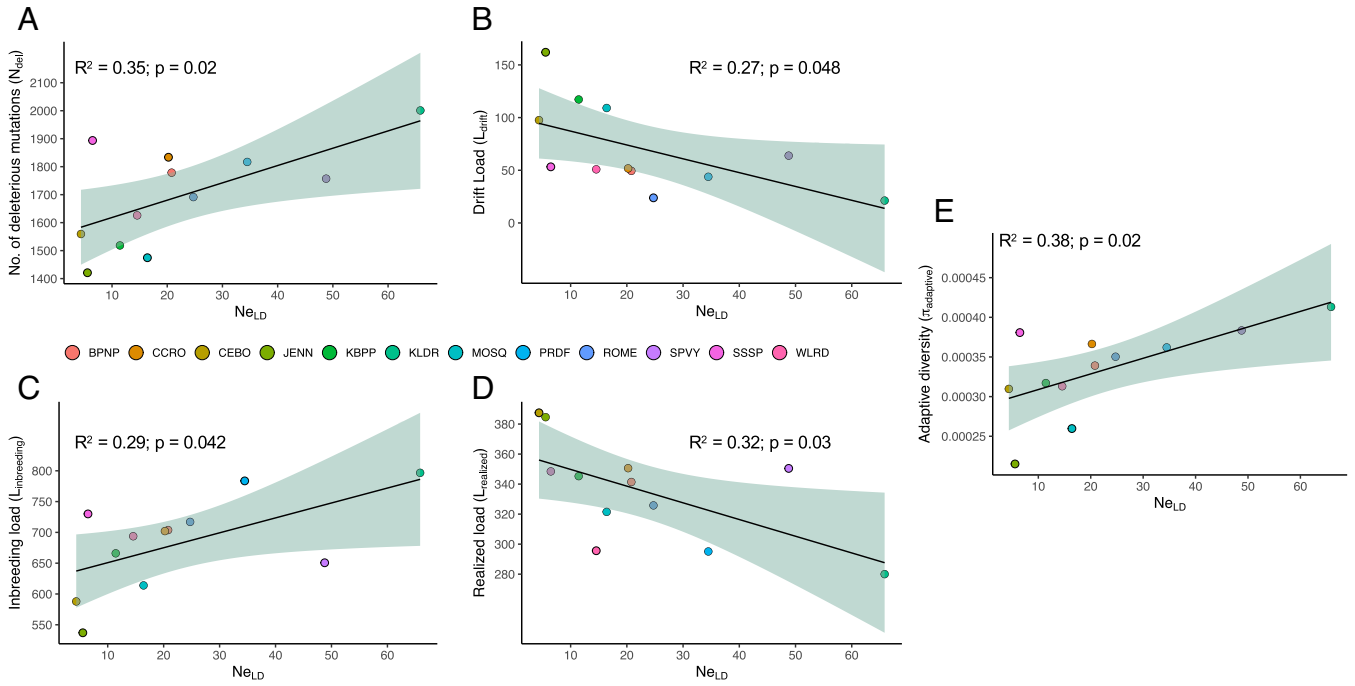

Fig. 3. Relationships between population estimates of functional and future neutral genetic diversity. For each population, functional genetic diversity was measured as (*A*–*D*) mutation load at a (*A* and *B*) population level or at an (*C* and *D*) individual level and as (*E*) adaptive nucleotide diversity ($\pi_{adaptive}$) at a population level. A future estimate of neutral genetic diversity was predicted by the effective population size estimated from linkage disequilibrium (Ne_{LD}). See *Materials and Methods* and *SI Appendix*, Table S4 for how each variable was calculated from *S. catenatus* population genomic data. The color-coded population origin for each data point is given in the legend. The black line represents the linear regression model, with the shaded region representing 95% CI. R^2 indicates the explained variance, and the *P*-value represents the statistical significance of the slope of the linear model.

effect sizes of $\pi_{geneDes}$ and Ne_{LD} using standardized regression coefficients (*SI Appendix*, Fig. S3). The results confirm that $\pi_{geneDes}$ is a better predictor of adaptive diversity, population-level load, and individual inbreeding load, whereas Ne_{LD}, which essentially is a measure of expected increase in inbreeding, is better at predicting individual realized load (*SI Appendix*, Fig. S3).

Discussion

Our findings support the long-standing but recently controversial claim that neutral genetic diversity can predict genome-wide levels of positive and negative functional diversity in small populations of threatened species (14, 30). This reinforces the idea that neutral diversity can be used to indirectly assess the amount of genetic erosion in small populations without the need to directly measure functional diversity at the population (51) or species level (8). This work also demonstrates the value of whole genome resequencing for conservation actions by providing a template for parsing genomic data into neutral and functional components, which can then be used to assess the relationship between neutral and functional variation in other species of conservation concern (36, 51).

Our results also provide nuance to the neutral-functional diversity controversy by showing that the ability of neutral genetic variation to predict levels of functional variation depends on the time scale of the evolutionary processes that generate these correlations. Specifically, historical measures of neutral diversity in the form of population-level nucleotide diversity (π) were substantially better at predicting levels of different components of mutation load and adaptive diversity than future estimates of diversity at equilibrium inferred from recent population size. At the population level, values of nucleotide diversity are predicted to reach equilibrium over a timescale of roughly $4N_e$ generations (31), which, in Massasauga rattlesnakes, reflects a time scale of >1,000s ybp, or well before major human impacts (48). In contrast, Ne_{LD} (52) estimates effective population size over the past 1 to 2 generations (53) and so is a better predictor of the expected levels of neutral variation that should exist at equilibrium in these populations at drift-migration equilibrium in the future (32, 33). We acknowledge that is also possible that the reduced correlations between levels of functional genetic diversity and neutral diversity at contemporary timescales are simply a by-product of the weak but significant positive correlations between population values for π and Ne_{LD} (see above) and that neutral variation may no longer be useful in predicting functional variation in present-day or future populations. Finally, we note that near-pervasive absence of mutation-drift-selection-migration equilibrium in wild populations means that contemporary census and effective population sizes are not expected to be strongly correlated with levels of neutral and functional genetic diversity due to stochastic variation in contemporary effective size. However, contemporary N_e is still highly useful in the context of conservation biology because populations with small N_e will be subject to rapid genetic erosion in the future if there is no increase in population size or increased immigration.

We interpret the weakening of the correlation between future measures of neutral and functional diversity in these snakes as reflecting a nonequilibrium state that has resulted from recent declines in population sizes (48) that have yet to be tracked by positive and negative levels of functional diversity in individual populations (54). This situation arises because historical diversity-based estimates may reflect the "ghost of evolution past" (55) in that the observed correlations between historical estimates of neutral and functional diversity are present but do not reflect expected correlations in future populations after they regain genetic equilibrium. These results were also demonstrated using forward-time population genetic model simulations in a study that showed that demographic events

such as a population bottleneck can lead to a breakdown in correlations between measures of neutral and functional diversity and can take many generations to be reestablished while populations return to genetic equilibrium (30).

Other genetic diversity studies in these snake populations have provided evidence that they are not currently in equilibrium with respect to levels of neutral genetic diversity relative to effective population size. For example, the levels of neutral diversity estimated from RAD loci in these populations were up to 50% higher than the equilibrium levels of contemporary N_e predicted by demographic models (48). The models suggested that if population sizes remained at present levels, then populations would experience substantial declines in genome-wide neutral diversity over the next 100 y. Likewise, persistence of functional variation within coding sequences of snake venom proteins in these populations despite small N_e (49) is consistent with the idea that much of the adaptive variation present in populations may represent an example of "drift debt," a nonequilibrium state where present-day levels of variation overestimate the amount of functional genetic diversity present in future populations (55) (although it is also possible that it is being maintained by strong selection despite small contemporary effective population sizes). Such populations may be poised to enter an extinction vortex, with the true genetic cost of living at their current population size yet to be realized (49).

Our results suggest that significant correlations may exist between measures of neutral and functional diversity in small populations of threatened species, but these may be of limited usefulness for conservation planning because they reflect past but not present or potential future conditions due to nonequilibrium demography. This circumstance is an example of the gap that can exist between the historical timescales over which the evolutionary processes relevant to conservation genomics operate and the contemporary timescales of ecological processes that are most relevant to conservation biologists dealing with rapid recent population declines in the Anthropocene (18). Conservation genetics is challenging because populations of threatened species face constantly changing environments that alter key characteristics such as population size, making predictions from genomic data difficult.

Our results rest on several important assumptions. The first is we assume that nonsynonymous variants in genes under strong positive and negative selection reflect plausible measures of mutation load and adaptive variation. However, compensatory mutations may modify the effect of substitutions our scheme classifies as either deleterious mutations or adaptive variants (see ref. 8), and factors such as dominance, epistasis, pleiotropy, and purging may also complicate the relationship between mutation load, adaptive diversity and fitness (e.g., ref. 24). Geographic variation in habitat can also alter the impact of deleterious or adaptive mutations among populations (56). It also seems unlikely that our approach detects all adaptive variants that could arise under different types of selection. Because we have focused on measuring variants that increase levels of functional polymorphism in populations, adaptive variants that evolve through other evolutionary scenarios (e.g., frequency dependent selection) should also be assessed. For these reasons, the impact of the higher mutation load and lower levels of adaptive variation in smaller populations will be challenging to quantify in the absence of direct fitness data, such as individual reproductive success and survival. We note that the degree to which selection acts to maintain or eliminate specific variants could be examined through analyses that combine modeling to account for the confounding effects of demography on detecting selection with methods that detect selection within and between populations (57) as was recently conducted to detect selection on venom genes in rattlesnakes (58).

Second, our use of nonsynonymous variants in genes under strong positive and negative selection makes it likely that we are only considering a set of large effect variants that underlie genetic load and adaptive diversity and are failing to consider other loci with smaller yet substantial impacts on fitness. These large-effect "golden SNPs" (cf. ref. 10) represent a subset of variants with fitness effects that are important to focus on for conservation genetic efforts, but we run the risks of incorporating the flaws of gene-targeted conservation analyses (59). However, our approach of using polymorphism and divergence data provides an agnostic method to identify genes under strong positive and negative selection, which represent a sizeable subset of functional variants that are potentially important contributors to fitness in these snakes.

Finally, it is possible that despite a lack of phylogeographic structure in this species (45) different populations may share different amounts of evolutionary history, which would introduce nonindependence into our analyses of estimates of diversity from each population. However, without a more detailed knowledge of the demographic history of these populations (for example, how and when they were connected by migration), it is difficult to determine how best to correct for this nonindependence. Future work will combine this type of demographic inference with models that incorporate the interpopulation relatedness structure into our analyses of the predictors of adaptive and deleterious variation across populations.

Materials and Methods

***S. catenatus* Genome Sequencing and Assembly.** Details of the methods used to generate a reference chromosome-level genome assembly are provided in *SI Appendix*. Briefly, DNA from two female *S. catenatus* individuals was sequenced using PacBio® and Illumina® methodologies, and a hybrid assembly was generated using the MaSuRCA v.3.4.1 genome assembler (60). To improve assembly length, we generated Hi-C sequencing data for genome scaffolding through Dovetail Genomics. The resultant genome was annotated using the Maker pipeline and following established protocols (61, 62). The final assembly contained 2,004 scaffolds, with a scaffold N50 of 195.71 MB, a contig N50 of 0.81 MB, and 18 chromosome-scale scaffolds.

Whole Genome Population Resequencing. To estimate functional and neutral genetic diversity in *S. catenatus* populations, we analyzed whole genome data from 12 different previously studied geographic locations (designated populations; *SI Appendix*, Table S3 and Dataset S3) (26, 48, 49). We combined previously sequenced whole genome data (BioProject PRJNA750087; BioSample SAMN20438760–859) with new individuals sequenced for this study (N = 117; Dataset S3). DNA extraction and Illumina® paired-end 150-bp reads sequencing were performed as described above (*SI Appendix, Supplementary Methods*). We also included 10 previously sequenced whole genomes from sister species, *S. tergeminus*, as an outgroup (26). All sequences were mapped to the *S. catenatus* Hi-C reference genome with BWA v.0.7.17 (63) using the mem algorithm. Sequence preprocessing and mapping was performed following GATK "Best Practice Workflow" (64) (*SI Appendix, Supplementary Methods*). After read mapping, we removed any individual with mean depth of coverage <10× or if the proportion of the reference genome covered at 10× was <40%. Ultimately, we analyzed 110 *S. catenatus* genomes with mean 14.2× (10.0× to 27.3×) (min to max) depth of coverage sequenced at mean 74.1% (47.2 to 95.2%) (min to max) of the reference genome covered at 10× depth. We also analyzed 10 *S. tergeminus* samples with 14.1× (10.9× to 18.3×) depth and 71.2% (56.7 to 87.7%) of reference genome covered at 10×.

Genotype Calling, Variant Filtration, and Variant Annotation. We identified biallelic SNP variants and called individual genotypes using GATK v4.1.2.0 (65) (*SI Appendix, Supplementary Methods*). We used "QD < 2.0; MQ < 20.0; |MQRankSum| > 3.0; |ReadPosRankSum| > 3.0; SOR > 5.0" to filter low-quality SNPs (24). SNPs with genotypes missing in >20% of samples or SNPs with MAF < 0.05 were filtered out. Our genotype calling and filtering pipeline resulted in

8,831,200 SNPs. We used the reference genome annotation file (.gff) and reference genome (.fasta) to functionally annotate genome-wide SNPs using SnpEff v.4.3 (66). Out of ~8.8 million SNPs, 80% of SNPs were in intergenic regions, whereas only 406,098 SNPs (~1.4%) were in exonic regions. Variant partitioning by different genomic regions, variant statistics, and diversity estimates were calculated using VCFTools (67).

Identifying Functional Genomic Variation. To identify SNP variants that plausibly represent different components of functional genetic diversity, we assumed that a) nonsynonymous variants in genes that were under strong positive selection during divergence between sister taxa also represent adaptive mutations within each population, and b) nonsynonymous variants in genes under strong negative (purifying) selection will represent deleterious mutations. This assumption is based on the expectation that the evolutionary history of selection on a gene at the interspecific level will also reflect the type of selection acting at the intraspecific level (68, 69) as recently demonstrated for some mammals (70). We estimated the type and strength of selection acting on all annotated protein-coding genes by calculating the modified Neutrality Index called DoS (35) using polymorphism and divergence data from *S. catenatus* and *S. tergeminus*. For a gene, DoS can be estimated as:

$$DoS = \frac{D_n}{D_n + D_s} - \frac{P_n}{P_n + P_s},$$

where D_n is the number of nonsynonymous substitutions between *S. catenatus* and *S. tergeminus*, D_s is the number of synonymous substitutions between *S. catenatus* and *S. tergeminus* within the gene, P_n is the number of nonsynonymous polymorphisms within *S. catenatus*, and P_s is the number of synonymous polymorphisms within *S. catenatus*.

To minimize drift effects on intraspecific polymorphism data, we used polymorphism data from the KLDR population of *S. catenatus* (*SI Appendix*, Fig. S2 and Table S2) which has a relatively large effective population size (48) and a reference population of the outbred sister species, *S. tergeminus* from Cheyenne Bottoms (*SI Appendix*, Fig. S2 and Table S2) to estimate divergence (26). We removed genes with undefined DoS and further restricted our analyses to genes with $P_n + D_n \geq 4$ which are potentially informative about selection (41, 42). We use this approach only to identify loci under different types of selection and not to evaluate per se whether specific genetic variants were functionally adaptive or deleterious. Our approach focuses on a small subset of the genetic variants that are potentially deleterious or adaptive namely those in sequences of protein-coding genes. For practical reasons, we do not attempt to catalog regulatory variants that influence gene expression and can clearly play a role in adaptive variation (e.g., ref. 71).

To estimate functional diversity, we designated nonsynonymous substitutions in protein-coding regions of the genes with the highest 10% of DoS values as possible adaptive variants while nonsynonymous substitutions in genes with the lowest 10% of DoS values represent deleterious mutations. Analysis that used more extreme categorizations of loci (upper and lower 5% of DoS values) yielded similar results (results not shown) and so we used the 10% cutoff because it represents a more comprehensive survey of functional variation in the *S. catenatus* genome. We also designated SNPs that were identified by SnpEff to cause LoF in coding sequences as deleterious mutations because LoFs are likely under strong purifying selection as shown by their negative impacts on fitness (50). We then used these polymorphism data to generate empirical estimates of negative and positive functional diversity at population and individual levels (*SI Appendix*, Table S4 for specific metrics used and *SI Appendix, Supplementary Methods* for additional details).

Negative functional diversity is usually equated to different components of genetic load (40) yet this requires estimates of the fitness impacts of specific mutations that were unavailable to us. As a proxy, we used estimates of mutation load defined in terms of the frequency and number of putatively deleterious mutations. Following (30) we estimated two components of load at the population and individual levels: mutation inbreeding load and mutation drift load. At the population level, both measures of load depend on both the absolute number and frequency of deleterious mutations that are segregating within the population. A population with a greater number of deleterious mutations will have higher mutation inbreeding load (and thus presumably a greater negative impact on population fitness). Similarly, a population where drift and weaker selection has led to an increased allele frequency of many deleterious mutations to near fixation (MAF > 0.9) would have higher mutation drift load. For each population, we estimated the number of deleterious mutations that are segregating (MAF > 0) within the population (N_{del}) and drift load (L_{drift}) as the number of deleterious mutations that are nearly fixed within the population (MAF > 0.9). Mean and SD for N_{del} and L_{drift} were derived from 100 resamplings of 5 individuals from each population (lowest sample size in our dataset; *SI Appendix*, Table S2) to account for differences in sample sizes between populations.

At the individual level, we estimated mutation inbreeding load and mutation drift load by partitioning the two components of individual load by the zygosity of deleterious alleles. $L_{inbreed}$ is the total number of deleterious SNPs that are present as heterozygotes whereas $L_{realized}$ is the total number of SNPs that are homozygous for the deleterious allele within each individual genome (23, 40). The population mean and SD for each of $L_{inbreed}$ and $L_{realized}$ were then estimated from individual estimates within each population.

Finally, at the population level, we measured positive functional diversity in terms of adaptive diversity ($\pi_{adaptive}$) which is a measure of how much functional (i.e., nonsynonymous) diversity exists in genes under strong positive selection in a given population. We used this measure as a proxy for the additive diversity at a quantitative trait where diversity impacts the response of the population to present or future selection. For each population, $\pi_{adaptive}$ is the pairwise nucleotide diversity at the nonsynonymous SNPs within the protein-coding regions of genes in the upper 10% of the DoS distribution. This measure accounts for variation due to the number of variable sites within coding region and for differences in among populations in variation in allele frequency distributions. We counted all protein-coding (variant and invariant) sequences in the total number of sites for π adaptive calculations (*SI Appendix, Supplementary Methods*). We did not calculate an individual-based estimate of adaptive diversity.

Estimates of Neutral Diversity. Our goal was to examine the relationship between direct and indirect estimates of neutral diversity and functional diversity at population and individual levels that reflect levels of neutral diversity at historical- and contemporary-time scales. For the estimates of neutral diversity, we used variation at putatively neutral markers from genome-wide data. We defined neutral regions as genomic regions that exist within "gene deserts," i.e., at least 500 kb upstream and downstream from an annotated gene (gene deserts) (72) (*SI Appendix, Supplementary Methods*). For direct estimates of historical genetic diversity, we used mean pairwise nucleotide diversity in 1-kb windows of gene desert regions of a given population ($\pi_{geneDes}$). Because $\pi_{geneDes}$ is based on levels of standing neutral variation, which is impacted by both effective size and mutation (73), it primarily reflects the intensity of genetic drift experienced over many generations (34). It is probable that a few noncoding mutations within gene desert regions may be influenced by natural selection, but such mutations are rare and would not likely affect our mean estimates of genome-wide neutral diversity (74). As a proxy for future levels of neutral diversity in populations at drift-mutation equilibrium, we estimated contemporary effective population size using the degree of population-level linkage disequilibrium (Ne_{LD}) between physically unlinked SNPs within gene desert regions (52) (*SI Appendix, Supplementary Methods*). This method estimates contemporary effective size over a time scale of 1 to 2 generations (<5 y–see ref. 48). We assume that contemporary effective size is closely related to future equilibrium levels of neutral diversity once populations reach drift mutation equilibrium (see discussion in ref. 47).

Correlations between Functional and Neutral Genetic Diversity. Our study focuses on testing predictions of the relationships between functional and neutral genetic diversity at a population level by fitting linear regression models similar to ref. 30, yet some estimates of functional diversity (mutation load) were also calculated on a per-individual level. As pointed out by ref. 40, the availability of whole genomes for individuals opens the door for analyses that can account for individual variation around the population mean. To explore whether estimating mutation load at the individual level ($L_{inbreed}$ and $L_{realized}$) yields similar results, we modeled the genetic load versus neutral genetic diversity relationships using mixed-effect models which allowed us to account for intrapopulation variation as a random effect nested within individuals' population assignment. All models were run using the lme function of the nlme package in R (75). To assess model fit, we inferred R^2 values for each model using the r.squaredGLMM function of the MuMIn package (76).

Data, Materials, and Software Availability. Raw sequence data used in this study are available in NCBI's Short Read Archive BioProject Accession No. PRJNA975611 (Submission metadata is provided in Dataset S4) (77). The *S. catenatus* reference genome and annotation is available on GenBank under Accession No. JASPKD000000000 (78) and on Dryad (https://doi.org/10.5061/dryad.r4xgxd2k2) (79). The scripts developed for data analysis and visualization can be publicly accessed on GitHub at https://github.com/samarth8392/FGDvNGD (80).

ACKNOWLEDGMENTS. We thank all those individuals who have provided samples or assisted with collections across the range of *S. catenatus* over the past 25 y–this work would not be possible without their help, and we are deeply appreciative. Many individuals assisted, but we especially thank Jeff Davis, Michael Dreslik, Brian Fedorko, Tony Frazier, Kim Frolich, Dan Harvey, Matt Kowalski, Greg Lipps, Scott Martin, 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. Marty Kardos provided detailed comments on the manuscript at all stages which significantly improved the paper. We also thank J. Andrew DeWoody, Laura Kubatko, and Matt Hahn for discussion, members of the Bradburd and Gibbs Labs for comments on early versions of the manuscript, and Kendra Wecker, Kate Parsons, and Carolyn Caldwell from the Ohio Division of Wildlife for their long-standing support of conservation genetics work by H.L.G. on endangered snakes. 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 and funds from Michigan State University. H.L.G. was supported by NSF (USA) Grant DEB 1638872 during the preparation of the manuscript. Research reported in this publication was also supported by the National Institute of General Medical Sciences of the NIH under Award Number R35GM137919 (awarded to G.S. B.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Computational analyses were performed using resources provided by the Ohio Supercomputer Center.

Author affiliations: [a]Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, OH 48824; [b]Ohio Biodiversity Conservation Partnership, The Ohio State University, Columbus, OH 43210; [c]Evolution and Behavior Program, Department of Integrative Biology, Ecology, Michigan State University, East Lansing, MI 48824; and [d]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109

1. R. Frankham *et al.*, *Genetic Management of Fragmented Animal and Plant Populations* (Oxford University Press, 2017), 10.1093/oso/9780198783398.001.0001.
2. G. Leroy *et al.*, Next-generation metrics for monitoring genetic erosion within populations of conservation concern. *Evol. Appl.* **11**, 1066–1083 (2017).
3. C. van Oosterhout, Mutation load is the spectre of species conservation. *Nat. Ecol. Evol.* **4**, 1004–1006 (2020).
4. R. Bijlsma, V. Loeschcke, Genetic erosion impedes adaptive responses to stressful environments. *Evol. Appl.* **5**, 117–129 (2012).
5. G. Caughley, Directions in conservation biology. *J. Anim. Ecol.* **63**, 215–244 (1994).
6. P. W. Hedrick, M. E. Gilpin, "Genetic effective size of a metapopulation" in *Metapopulation Biology*, I. Hanski, M.E. Gilpin, Eds. (Academic Press, 1997), 10.1016/b978-012323445-2/50011-0, pp. 165–181.
7. K. Ralls, P. Sunnucks, R. C. Lacy, R. Frankham, Genetic rescue: A critique of the evidence supports maximizing genetic diversity rather than minimizing the introduction of putatively harmful genetic variation. *Biol. Conserv.* **251**, 108784 (2020).
8. A. P. Wilder *et al.*, The contribution of historical processes to contemporary extinction risk in placental mammals. *Science* **380**, eabn5856 (2023).
9. R. D. H. Barrett, H. E. Hoekstra, Molecular spandrels: Tests of adaptation at the genetic level. *Nat. Rev. Genet.* **12**, 767–780 (2011).
10. M. V. Rockman, The Qtn program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution* **66**, 1–17 (2012).
11. C. Bozzuto, I. Biebach, S. Muff, A. R. Ives, L. F. Keller, Inbreeding reduces long-term growth of Alpine ibex populations. *Nat. Ecol. Evol.* **3**, 1359–1364 (2019).
12. R. Lande, S. Shannon, The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* **50**, 434–437 (1996).
13. I. Saccheri *et al.*, Inbreeding and extinction in a butterfly metapopulation. *Nature* **392**, 491–494 (1998).
14. J. A. DeWoody, A. M. Harder, S. Mathur, J. R. Willoughby, The long-standing significance of genetic diversity in conservation. *Mol. Ecol.* **30**, 4147–4154 (2021).
15. L. N. Carley, W. F. Morris, R. Walsh, D. Riebe, T. Mitchell-Olds, Are genetic variation and demographic performance linked? *Evol. Appl.* **15**, 1888–1906 (2022), 10.1111/eva.13487.
16. S. W. Fitzpatrick *et al.*, Genomic and fitness consequences of genetic rescue in wild populations. *Curr. Biol.* **30**, 517–522.e5 (2020).
17. D. A. G. A. Hunt, J. D. DiBattista, A. P. Hendry, Effects of insularity on genetic diversity within and among natural populations. *Ecol. Evol.* **12**, e8887 (2022).
18. J. C. Teixeira, C. D. Huber, The inflated significance of neutral genetic diversity in conservation genetics. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015096118 (2021).
19. J. R. Chapman, S. Nakagawa, D. W. Coltman, J. Slate, B. C. Sheldon, A quantitative review of heterozygosity-fitness correlations in animal populations. *Mol. Ecol.* **18**, 2746–2765 (2009).
20. D. H. Reed, R. Frankham, How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* **55**, 1095–1103 (2001).
21. P. A. Morin *et al.*, Reference genome and demographic history of the most endangered marine mammal, the vaquita. *Mol. Ecol. Res.* **21**, 1008–1020 (2020).
22. J. A. Robinson *et al.*, Genomic flatlining in the endangered island fox. *Curr. Biol.* **26**, 1183–1189 (2016).
23. S. Mathur, J. A. DeWoody, Genetic load has potential in large populations but is realized in small inbred populations. *Evol. Appl.* **14**, 1540–1557 (2021).
24. C. Grossen, F. Guillaume, L. F. Keller, D. Croll, Purging of highly deleterious mutations through severe bottlenecks in Alpine ibex. *Nat. Commun.* **11**, 1001 (2020).
25. S. Mathur, J. M. Tomeček, L. A. Tarango-Arámbula, R. M. Perez, J. A. DeWoody, An evolutionary perspective on genetic load in small, isolated populations as informed by whole genome resequencing and forward-time simulations. *Evolution* **77**, 690–704 (2023).
26. A. Ochoa, H. L. Gibbs, Genomic signatures of inbreeding and mutation load in a threatened rattlesnake. *Mol. Ecol.* **30**, 5454–5469 (2021).
27. J. B. S. Haldane, The effect of variation of fitness. *Am. Nat.* **71**, 337–349 (1937).
28. R. J. Baker, G. A. Mengden, J. J. Bull, Karyotypic studies of thirty-eight species of North American snakes. *Copeia* **1972**, 257–265 (1972).
29. D. A. Bell *et al.*, The ecological causes and consequences of hard and soft selection. *Ecol. Lett.* **24**, 1505–1521 (2021).
30. M. Kardos *et al.*, The crucial role of genome-wide genetic variation in conservation. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104642118 (2021).
31. M. P. Hare *et al.*, Understanding and estimating effective population size for practical application in marine species management. *Conserv. Biol.* **25**, 438–449 (2011).
32. J. Wakeley, O. Sargsyan, Extensions of the coalescent effective population size. *Genetics* **181**, 341–345 (2009).
33. R. S. Waples, C. Do, Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evol. Appl.* **3**, 244–262 (2010).
34. K. Nadachowska-Brzyska, M. Konczal, W. Babik, Navigating the temporal continuum of effective population size. *Methods Ecol. Evol.* **13**, 22–41 (2021).
35. N. Stoletzki, A. Eyre-Walker, Estimation of the neutrality index. *Mol. Biol. Evol.* **28**, 63–70 (2010).
36. P. A. Hohenlohe, W. C. Funk, O. P. Rajora, Population genomics for wildlife conservation and management. *Mol. Ecol.* **30**, 62–82 (2020).
37. N. Galtier, L. Duret, Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution *Trends Genet.* **23**, 273–277 (2007).
38. J. D. Jensen, K. R. Thornton, C. D. Bustamante, C. F. Aquadro, On the utility of linkage disequilibrium as a statistic for identifying targets of positive selection in nonequilibrium populations. *Genetics* **176**, 2371–2379 (2007).
39. H. A. Orr, The population genetics of beneficial mutations. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 1195–1201 (2010).
40. G. Bertorelle *et al.*, Genetic load: Genomic estimates and applications in non-model animals. *Nat. Rev. Genet.* **23**, 492–503 (2022).
41. C. D. Bustamante *et al.*, Natural selection on protein-coding genes in the human genome. *Nature* **437**, 1153–1157 (2005).
42. Y. F. Li, J. C. Costello, A. K. Holloway, M. W. Hahn, "Reverse ecology" and the power of population genomics. *Evolution* **62**, 2984–2994 (2008).
43. US Fish and Wildlife Service, Final Rule: Endangered and threatened wildlife and plants; Threatened species status for the eastern massasauga rattlesnake. *Fed. Regis.* **81**, 67193–67214 (2016).
44. Government of Canada, Species at risk public registry (2009). https://www.sararegistry.gc.ca. Accessed 15 February 2023.
45. M. G. Sovic, A. C. Fries, H. L. Gibbs, Origin of a cryptic lineage in a threatened reptile through isolation and historical hybridization. *Heredity* **117**, 358–366 (2016).
46. J. E. Chiucchi, H. L. Gibbs, Similarity of contemporary and historical gene flow among highly fragmented populations of an endangered rattlesnake. *Mol. Ecol.* **19**, 5345–5358 (2010).
47. S. A. Martin, G. J. Lipps, H. L. Gibbs, Pedigree-based assessment of recent population connectivity in a threatened rattlesnake. *Mol. Ecol. Res.* **21**, 1820–1832 (2021).
48. M. Sovic, A. Fries, S. A. Martin, H. L. Gibbs, Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, Sistrurus catenatus. *Evol. Appl.* **12**, 664–678 (2019).
49. A. Ochoa *et al.*, Drift, selection and adaptive variation in small populations of a threatened rattlesnake. *Mol. Ecol.* **29**, 2612–2625 (2020).
50. J. A. Robinson *et al.*, Genomic signatures of extensive inbreeding in Isle Royale wolves, a population on the threshold of extinction. *Sci. Adv.* **5**, eaau0757 (2019).
51. J. A. DeWoody *et al.*, The Threatened Species Imperative: Conservation assessments would benefit from population genomic insights. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2210685119 (2022).
52. R. S. Waples, A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conserv. Genet.* **7**, 167–184 (2006).
53. R. S. Waples, T. Antao, G. Luikart, Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics* **197**, 769–780 (2014).
54. L. Faust, J. Szymanski, M. Redmer, Range wide extinction risk modeling for the eastern massasauga rattlesnake (Sistrurus catenatus catenatus) (Alexander Center for Applied Population Biology, Lincoln Park Zoo, Chicago, IL, 2011), 10.3996/052017-JFWM-041. S1. (see Supplemental Material, Reference S1).

55. D. L. Gilroy, K. P. Phillips, D. S. Richardson, C. van Oosterhout, Toll-like receptor variation in the bottlenecked population of the Seychelles warbler: Computer simulations see the 'ghost of selection past' and quantify the 'drift debt'. *J. Evol. Biol.* **30**, 1276–1287 (2017).
56. J. A. Mee, S. Yeaman, Unpacking conditional neutrality: Genomic signatures of selection on conditionally beneficial and conditionally deleterious mutations. *Am. Nat.* **194**, 529–540 (2019).
57. P. Johri *et al.*, Recommendations for improving statistical inference in population genomics. *PLOS Biol.* **20**, e3001669 (2022).
58. D. R. Schield *et al.*, The roles of balancing selection and recombination in the evolution of rattlesnake venom. *Nat. Ecol. Evol.* **6**, 1367–1380 (2022).
59. M. Kardos, A. B. A. Shafer, The peril of gene-targeted conservation. *Trends in Ecol. Evol.* **33**, 827–839 (2018).
60. A. V. Zimin *et al.*, The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
61. D. R. Schield *et al.*, The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. *Genome Res.* **29**, 590–601 (2019).
62. M. J. Margres *et al.*, The Tiger Rattlesnake genome reveals a complex genotype underlying a simple venom phenotype. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014634118 (2021).
63. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
64. G. A. Auwera *et al.*, From fastq data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 11.10.1–11.10.33 (2018).
65. A. McKenna *et al.*, The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
66. P. Cingolani *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* **6**, 80–92 (2014).
67. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
68. J. H. McDonald, M. Kreitman, Adaptive protein evolution at the Adh locus in Drosophila. *Nature* **351**, 652–654 (1991).
69. S. A. Sawyer, D. L. Hartl, Population genetics of polymorphism and divergence. *Genetics* **132**, 1161–1176 (1992).
70. T. Latrille, N. Rodrigue, N. Lartillot, Genes and sites under adaptation at the phylogenetic scale also exhibit adaptation at the population-genetic scale. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2214977120 (2023).
71. A. J. Mason *et al.*, Venom gene sequence diversity and expression jointly shape diet adaptation in pitvipers. *Mol. Biol. Evol.* **39**, msac082 (2022).
72. S. Mathur, J. M. Tomeček, A. Heniff, R. Luna, J. A. DeWoody, Evidence of genetic erosion in a peripheral population of a North American game bird: The Montezuma quail (*Cyrtonyx montezumae*). *Conserv. Genet.* **20**, 1369–1381 (2019).
73. M. Nei, *Molecular Population Genetics and Evolution (Frontiers of Biology)* (North-Holland Publishing Comapany; American Elsevier Publishing Company, Amsterdam New York, 1975), p. xiii, p. 288.
74. R. T. Hietpas, J. D. Jensen, D. N. A. Bolon, Experimental illumination of a fitness landscape. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7896–7901 (2011).
75. J. Pinheiro, D. Bates, nlme: Linear and Nonlinear Mixed Effects Models (Version 3.1-162, R package, 2023). http://CRAN.R-project/package=nlme. Accessed 15 February 2023.
76. K. Barton, M. K. Barton, MuMin: Multi-model inference (R package Version 1 47.5, 2023). http//CRAN.R-project.org/package=MuMin. Accessed 15 February 2023.
77. S. Mathur, H. L. Gibbs, Correlation between genome-wide functional and neutral genetic diversity in Eastern Massasauga rattlesnakes. NCBI Sequence Read Archive. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA975611. Deposited 29 September 2023.
78. A. Mason, H. L. Gibbs, Sistrurus catenatus genome for "Functional Genomic Diversity is Correlated with Neutral Genomic Diversity in Populations of an Endangered Rattlesnake." NCBI GenBank. https://www.ncbi.nlm.nih.gov/nuccore/JASPKD000000000. Deposited 29 September 2023.
79. A. Mason, H. L. Gibbs, Reference genome dataset for "Functional Genomic Diversity is Correlated with Neutral Genomic Diversity in Populations of an Endangered Rattlesnake." Dryad. https://datadryad.org/stash/dataset/doi:10.5061/dryad.r4xgxd2k2. Deposited 30 September 2023.
80. S. Mathur, A. Mason, Analysis code for "Functional Genomic Diversity is Correlated with Neutral Genomic Diversity in Populations of an Endangered Rattlesnake". GitHub. https://github.com/samarth8392/FGDvNGD. Deposited 29 September 2023.