Genomes of an endangered rattlesnake show that neutral genetic variation predicts adaptive genetic variation and genetic load
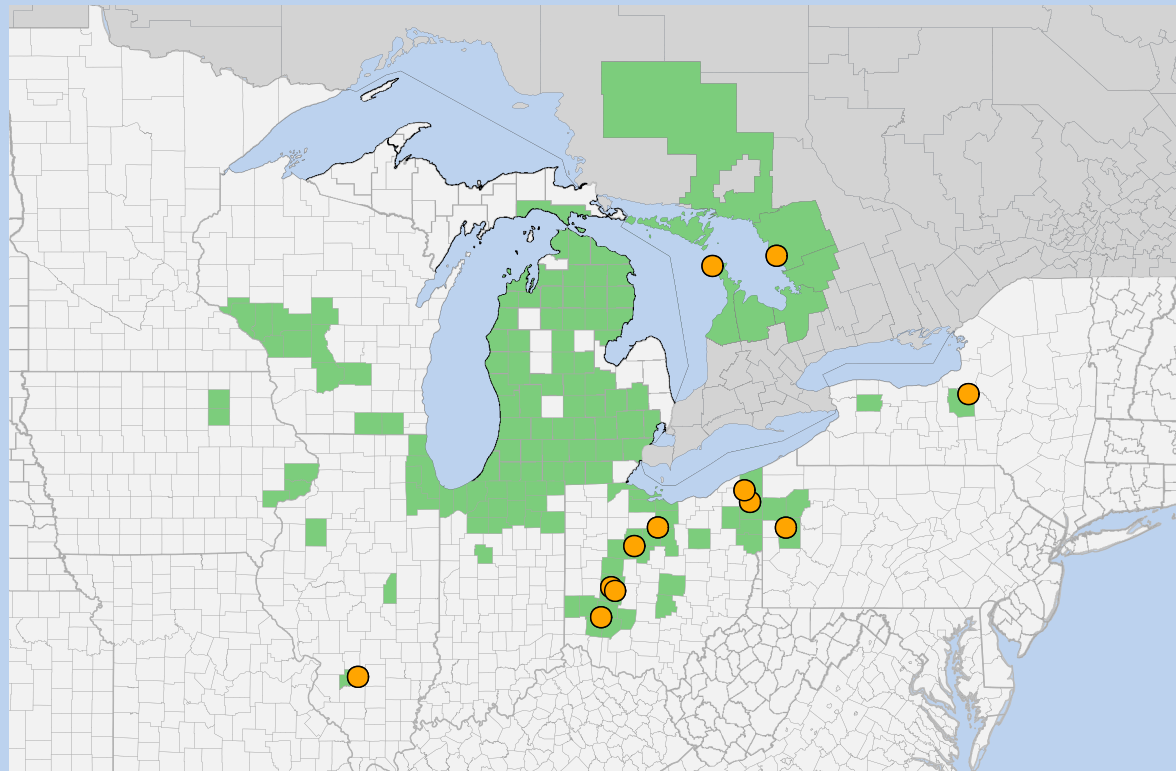
Marty Kardos[a,1] 🔾

Fig. 1. Eastern Massasauga rattlesnakes sampling sites (orange points) and spatial distribution. The Great Lakes and Atlantic Ocean are shown in blue. Canada is shown with dark gray background. The United States of America is shown with light gray background. Counties in the United States of America and census divisions in Canada with extant populations are shown in green. Locations of extant populations are from the US Fish and Wildlife Service (https://www.fws.gov/species/eastern-massasauga-sistrurus-catenatus) and Parks Canada (15).

One of the central tenets of conservation biology is that genetic variation is important to population viability. Small, isolated populations are expected to have lower genetic variation (e.g., heterozygosity), and more inbreeding (mating between relatives) than large, connected ones. Reduced genetic variation and higher inbreeding can reduce fitness (1). Numerous studies have demonstrated that inbreeding depression (lower fitness of more inbred individuals) can be strong (2) and reduce population growth (3–5). Most new mutations that affect protein structure are deleterious, and the constant input of new mutations means that all populations carry a "genetic load" (6). However, the genetic variation ultimately arising from mutation is essential for populations

Author affiliations: ᵃConservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA 98112

See companion article, "Functional genomic diversity is correlated with neutral genomic diversity in populations of an endangered rattlesnake," 10.1073/pnas.2303043120.

[1]Email: martin.kardos@noaa.gov.

https://doi.org/10.1073/pnas.2316880120 **1 of 3**

to adapt to changing ecological conditions (7). Avoiding problems associated with inbreeding depression and maintaining genetic variation and adaptive potential have long been viewed as key to conserving wild populations.

The genetic status of threatened populations has typically been assumed to be a function of genetic variation (e.g., heterozygosity) across the whole genome (8, 9). Populations with higher estimated heterozygosity are typically inferred to have lower inbreeding and higher adaptive potential. But what if we could focus instead only on the components of genetic variation that we think really matter: the functional loci that could impact fitness (10)? Whole genome sequencing makes it possible to measure genetic variation in coding genes and in other evolutionary-constrained genomic regions that might impact fitness (11). Would genetic variation just in these genomic regions provide improved measures of adaptive potential and genetic load than analyzing the whole genome, which includes regions that likely have little or no fitness effects?

> **Mathur et al.'s central finding was that adaptive and deleterious genetic variation were strongly correlated with π_{GeneDes}.**

The idea of focusing conservation assessments on functional genetic variation is appealing because it would mean that we could rapidly measure genetic status more efficiently than before by measuring only on the variation that matters. It has therefore been argued that conservation management should primarily focus on the functional genetic variation assumed to affect fitness (12, 13). Now that we can easily sequence entire genomes, an empirical assessment of the relationship between putatively neutral and functional variation is possible and could make significant headway in resolving the issue. In PNAS, Mathur et al. (14) accomplish this, using innovative analyses of genome sequence data from 12 populations of the endangered Massasauga rattlesnake (*Sistrurus catenatus*, Fig. 1).

The Massasauga rattlesnake was historically widespread in the Great Lakes region of North America, but human development has severely fragmented the species into small isolated populations (Fig. 1). Putatively neutral genetic variation, measured as nucleotide diversity in gene deserts (π_{GeneDes}), varied widely among the study populations, making for an ideal platform to test the theoretically predicted relationships between neutral and functional components of genetic variation. If molecular metrics of genetic load and adaptive potential measured in parts of the genome that show signatures of selection are unrelated to π_{GeneDes}, it would undermine the idea that genetic variation across the whole genome is a useful metric of fitness. Conversely, a strong relationship between π_{GeneDes} and functional variation would bolster the expectation of a strong correlation between neutral, deleterious, and adaptive variation (9).

Mathur et al.'s central finding was that adaptive and deleterious genetic variation were strongly correlated with π_{GeneDes}. Nucleotide diversity for coding regions putatively under positive selection was strongly correlated with π_{GeneDes}. Four measures of genetic load were also strongly correlated with π_{GeneDes}:

– N_{del}: the number of deleterious alleles in a population. More deleterious alleles imply a larger potential impact on population fitness
– $L_{\text{inbreeding}}$: the average number of heterozygous deleterious alleles, a measure of inbreeding load (the expected reduction in fitness with increasing inbreeding)
– L_{drift}: the number of high-frequency deleterious alleles, a measure of drift load (reduction in fitness due to the continuous fixation of deleterious alleles via genetic drift)
– L_{realized}: the average number of homozygous deleterious alleles, or "realized load." Because most deleterious alleles are partially recessive, their fitness effects are often only fully realized in homozygous genotypes.

Thus, Mathur et al.'s results suggest that neutral genetic diversity is indeed predictive of adaptive and deleterious genetic variation.

This finding is consistent with theoretical predictions (9) and is expected because each of the studied components of genetic load and adaptive potential is a function of heterozygosity. Additive genetic phenotypic variance (a metric of adaptive potential) increases with heterozygosity at quantitative trait loci (7). The inbreeding load increases with heterozygosity for deleterious alleles (16). Strong genetic drift reduces heterozygosity in smaller populations and renders natural selection incapable of preventing weakly deleterious alleles from drifting to fixation, resulting in an association between drift load and heterozygosity (17). It has long been known that heterozygosity is correlated across different parts of the genome when inbreeding varies among individuals (18) because inbreeding is equally likely to cause different genomic regions to be homozygous. The same applies to populations: the expected reduction in population heterozygosity due to genetic drift is the same genome-wide (19). Thus, Mathur et al.'s results support the long-standing expectation that neutral heterozygosity is a good predictor of different types of genetic load and adaptive potential.

Nevertheless, as in nearly all areas of biology, there are important nuances and exceptions to widely applicable rules of thumb. For example, the drift load can increase rapidly after a bottleneck, long before much genetic variation has been lost (9). Heritability, which is expected to decline after a bottleneck, can actually increase temporarily due to epistatic effects (20). Additionally, the fitness effects of inbreeding and lost genetic variation can depend heavily on population history and contemporary ecological context. Populations that have been small for a long time or have lost genetic variation due to slowly declining population size might have purged part of their genetic load via natural selection and might therefore be less susceptible to inbreeding depression than populations that crash rapidly. Populations with density-dependent fitness may also be less demographically susceptible to inbreeding depression because soft selection (where selection determines which, not how many individuals survive) tends to be dominant when fitness is density-dependent (6). Conversely, hard selection, which determines how many individuals survive and thus population growth, can dominate when fitness is density-independent (6). So, while the preponderance of the evidence shows that genetic variation is important to population

viability, this is not as simple as "higher genetic variation means increased population viability." Basic ecological details such as whether a population is currently declining, failing to recover despite sufficient habitat, and historical demography are well known to be crucial when considering conservation management actions such as genetic rescue (21).

There are important remaining questions regarding the relevance of putatively functional versus neutral genetic variation in conservation. First, we have very little understanding of the performance of population genetic and molecular approaches to identifying loci that are likely to affect fitness in contemporary ecological contexts. Even for species with substantial genetic data and accompanying functional assays, determining the fitness effects of any particular loci is extremely challenging. Population genetic methods, such as those used by Mathur et al., rely on patterns of interspecies divergence accumulated over very long periods of evolutionary history. An inevitable consequence of this historical approach is that many of the alleles identified as being positively selected may be either neutral or deleterious under current ecological conditions, and vice versa. A potential path to further evaluate the performance of molecular metrics of genetic load is to test whether individual fitness is better predicted by molecular measures of genetic load or by genome-wide homozygosity. This will require

distinguishing the effect sizes of two strongly correlated predictor variables (as shown by Mathur et al.), so very large sample sizes will likely be needed to have any chance of sufficient statistical power.

This uncertainty in the ability of molecular and population genetic methods to accurately identify deleterious and adaptive alleles implies that the loci that are important to fitness may never be completely knowable. Vrijenhoek and Leberg (22), in a now >30-y-old iteration of this same debate, argued that designing captive breeding programs specifically to preserve genetic variation only at loci thought to be under strong selection (13) amounted to "throwing the baby out with the bath water." In an argument that remains salient today (9), Vrijenhoek and Leberg (22) argued that we should aim to preserve genetic variation across the whole genome because the genetic variation that matters is mostly undescribed, and probably resides across the whole genome, including in noncoding regions with unknown function. Mathur et al.'s study makes an important contribution to this continuing debate by providing empirical support for the theoretical predictions which are the basis for the long-standing focus on the conservation of genetic variation across the whole genome in the field of conservation biology.

1.   S. Wright, Evolution in Mendelian populations. *Genetics* **16**, 97–159 (1931).
2.   D. Charlesworth, J. H. Willis, The genetics of inbreeding depression. *Nat. Rev. Genet.* **10**, 783–796 (2009).
3.   P. Leberg, Influence of genetic variability on population growth: Implications for conservation. *J. Fish Biol.* **37**, 193–195 (1990).
4.   J. Hogg, S. Forbes, B. Steele, G. Luikart, Genetic rescue of an insular population of large mammals. *Proc. R. Soc. B Biol. Sci.* **273**, 1491–1499 (2006).
5.   M. Kardos *et al.*, Inbreeding depression explains killer whale population dynamics. *Nat. Ecol. Evol.* **7**, 675–686 (2023).
6.   A. F. Agrawal, M. C. Whitlock, Mutation load: The fitness of individuals in populations where deleterious alleles are abundant. *Ann. Rev. Ecol. Evol. Syst.* **43**, 115–135 (2012).
7.   D. S. Falconer, T. F. C. Mackay, *Introduction to Quantitative Genetics* (Pearson, ed. 4, 1996).
8.   M. A. Cronin, R. Shideler, L. Waits, R. J. Nelson, Genetic variation and relatedness in grizzly bears in the Prudhoe Bay region and adjacent areas in northern Alaska. *Ursus* **16**, 70–84 (2005).
9.   M. Kardos *et al.*, The crucial role of genome-wide genetic variation in conservation. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104642118 (2021).
10.   G. Bertorelle *et al.*, Genetic load: Genomic estimates and applications in non-model animals. *Nat. Rev. Genet.* **23**, 492–503 (2022).
11.   J. A. Robinson *et al.*, The critically endangered vaquita is not doomed to extinction by inbreeding depression. *Science* **376**, 635–639 (2022).
12.   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).
13.   A. L. Hughes, MHC polymorphism and the design of captive breeding programs. *Conserv. Biol.* **5**, 249–251 (1991).
14.   S. Mathur, A. J. Mason, G. S. Bradburd, H. L. Gibbs, Functional genomic diversity is correlated with neutral genomic diversity in populations of an endangered rattlesnake. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2303043120 (2023).
15.   Parks Canada, "Recovery strategy for the Massasauga (*Sistrurus catenatus*) in Canada" in *Species at Risk Act Recovery Strategy Series* (Parks Canada, Ottawa, 2015), p ix + 37 pp.
16.   N. E. Morton, J. F. Crow, H. J. Muller, An estimate of the mutational damage in man from data on consanguineous marriages. *Proc. Natl. Acad. Sci. U.S.A.* **42**, 855–863 (1956).
17.   P. W. Hedrick, A. García-Dorado, Understanding inbreeding depression, purging, and genetic rescue. *Trends Ecol. Evol.* **31**, 940–952 (2016).
18.   M. Szulkin, N. Bierne, P. David, Heterozygosity-fitness correlations: A time for reappraisal. *Evolution* **64**, 1202–1217 (2010).
19.   J. F. Crow, M. Kimura, *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).
20.   C. J. Goodnight, Epistasis and the effect of founder events on the additive genetic variance. *Evolution* **42**, 441–454 (1988).
21.   A. R. Whiteley, S. W. Fitzpatrick, W. C. Funk, D. A. Tallmon, Genetic rescue to the rescue. *Trends Ecol. Evol.* **30**, 42–49 (2015).
22.   R. C. Vrijenhoek, P. L. Leberg, Let's not throw the baby out with the bathwater: A comment on management for MHC diversity in captive populations. *Conserv. Biol.* **5**, 252–254 (1991).