An evolutionary perspective on contemporary genetic load in threatened species to inform future conservation efforts

Samarth Mathur[1], John Tomeček[2], Luis Tarango-Arámbula[3], Robert Perez[4], and Andrew DeWoody[1]

[1]Purdue University
[2]Texas A&M University
[3]Colegio de Postgraduados Campus San Luis Potosi
[4]Texas Parks and Wildlife Department

June 29, 2021

Abstract

In theory, genomic erosion can be reduced in fragile "recipient" populations by translocating individuals from genetically diverse "donor" populations. However, recent simulation studies have argued that such translocations can, in principle, serve as a conduit for new deleterious mutations to enter recipient populations. A reduction in evolutionary fitness is associated with a higher load of deleterious mutations and thus, a better understanding of evolutionary processes driving the empirical distribution of deleterious mutations is crucial. Here, we show that genetic load is evolutionarily dynamic in nature and that demographic history greatly influences the distribution of deleterious mutations over time. Our analyses, based on both demographically explicit simulations as well as whole genome sequences of potential donor-recipient pairs of Montezuma Quail (Cyrtonyx montezumae) populations, indicate that all populations tend to lose deleterious mutations during bottlenecks, but that genetic purging is pronounced in smaller populations with stronger bottlenecks. Despite carrying relatively fewer deleterious mutations, we demonstrate how small, isolated populations are more likely to suffer inbreeding depression as deleterious mutations that escape purging are homogenized due to drift, inbreeding, and ineffective purifying selection. We apply a population genomics framework to showcase how the phylogeography and historical demography of a given species can enlighten genetic rescue efforts. Our data suggest that small, inbred populations should benefit the most when assisted gene flow stems from genetically diverse donor populations that have the lowest proportion of deleterious mutations.

An evolutionary perspective on contemporary genetic load in threatened species to inform future conservation efforts

Samarth Mathur[1+*], John. M. Tomeček[2], Luis A. Tarango-Arámbula[3], Robert M. Perez[4], J. Andrew DeWoody[1,5]

[1]*Department of Biological Sciences, Purdue University, 915 W. State St., West Lafayette, IN 47907, USA*

[2]*Department of Rangeland, Wildlife, and Fisheries Management, Texas A&M University, 534 John Kimbrough, College Station, TX 77843, USA*

[3]*Colegio de Postgraduados, Campus San Luis Potosí, Iturbide 73, Salinas de Hidalgo, San Luis Potosí 78600, Mexico*

[4]*Texas Parks and Wildlife Department, La Vernia, TX, 78121 USA*

[5]*Department of Forestry and Natural Resources, Purdue University, 715 W. State St., West Lafayette, IN 47907, USA*

[+] *Current address: Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, 318 W 12th Ave, Columbus, OH 43210, USA*

Running Title: Evolutionary history of genetic load

*Corresponding author:Samarth Mathur E-mail: mathur.112@osu.edu

ORCID IDs:

Samarth Mathur: 0000-0002-6446-5718

John M. Tomeček: 0000-0002-7494-283X

Luis A. Tarango-Arámbula: 0000-0002-7662-1319

J. Andrew DeWoody: 0000-0002-7315-5631

Abstract

In theory, genomic erosion can be reduced in fragile "recipient" populations by translocating individuals from genetically diverse "donor" populations. However, recent simulation studies have argued that such translocations can, in principle, serve as a conduit for new deleterious mutations to enter recipient populations. A reduction in evolutionary fitness is associated with a higher load of deleterious mutations and thus, a better understanding of evolutionary processes driving the empirical distribution of deleterious mutations is crucial. Here, we show that genetic load is evolutionarily dynamic in nature and that demographic history greatly influences the distribution of deleterious mutations over time. Our analyses, based on both demographically explicit simulations as well as whole genome sequences of potential donor-recipient pairs of Montezuma Quail (*Cyrtonyx montezumae*) populations, indicate that all populations tend to lose deleterious mutations during bottlenecks, but that genetic purging is pronounced in smaller populations with stronger bottlenecks. Despite carrying relatively fewer deleterious mutations, we demonstrate how small, isolated populations are more likely to suffer inbreeding depression as deleterious mutations that escape purging are homogenized due to drift, inbreeding, and ineffective purifying selection. We apply a population genomics framework to showcase how the phylogeography and historical demography of a given species can enlighten genetic rescue efforts. Our data suggest that small, inbred populations should benefit the most when assisted gene flow stems from genetically diverse donor populations that have the lowest proportion of deleterious mutations.

Keywords:

Evolutionary history; Demography; Functional variation; Deleterious mutations; Montezuma Quail; Ring Species

Introduction

Anthropogenic activities reduce and subdivide natural habitat (Fahrig, 2003), which produces isolated populations that often become smaller, less diverse, and less fit over time due to genetic erosion (Bijlsma & Loeschcke, 2012; Frankham, 2005). The preservation of genetic diversity and fitness is paramount for most modern conservation efforts (Ralls et al., 2018). Assisted gene flow via translocations ("genetic rescue") aims to increase mean fitness in small, isolated populations. Genetic rescue increases fitness by providing allelic variants necessary for adaptation and by reducing inbreeding depression (Barrett & Charlesworth, 1991; Whiteley, Fitzpatrick, Funk, & Tallmon, 2015) via increased heterozygosity and the masking of deleterious recessive alleles (Frankham, 2015; Frankham et al., 2017; Ralls, Sunnucks, Lacy, & Frankham, 2020). Unfortunately, conservation efforts are hindered by our limited understanding of the fitness effects of allelic variants (mutations) and how evolutionary processes like demography and selection shape contemporary gene pools in wild populations (Bell et al., 2019).

Recent evidence shows that risk of inbreeding depression and subsequent extinction probability of endangered species may also be dependent on ancient and recent evolutionary history (Benazzo et al., 2017; Grossen, Guillaume, Keller, & Croll, 2020; Kyriazis, Wayne, & Lohmueller, 2020; Robinson, Brown, Kim, Lohmueller,

& Wayne, 2018; Robinson et al., 2019). A reduction in fitness is associated with a higher load of deleterious mutations and thus, a better understanding of evolutionary processes driving the empirical distribution of deleterious mutations is crucial. Here, our goal is to jointly infer the role of selection and demography on the distribution and fitness impact of deleterious mutations in small vulnerable populations; and to create a genetic assessment-based vulnerability framework to guide prioritizing populations that may maximize the benefits of assisted gene flow or genetic rescue (Bell et al., 2019).

Clearly, the quantity, nature, and distribution of deleterious mutations contribute to fitness (Ellegren & Sheldon, 2008; Kimura, Maruyama, & Crow, 1963) and are relevant to conservation, but identifying deleterious mutations and quantifying their contribution to loss of fitness (i.e. genetic load) is not straightforward (van Oosterhout, 2020). Genetic load is a dynamic attribute of a population (Fu, Gittelman, Bamshad, & Akey, 2014) and the evolutionary mechanisms that maintain deleterious alleles in populations are distinct from those that maintain these alleles in individuals (Lohmueller, 2014). At a population level, genetic load is expected to increase in small populations where drift predominates and selection against deleterious mutations is weak (Kimura et al., 1963; Ohta, 1992). However, strong selection and bottlenecks can also purge strongly deleterious mutations through these same evolutionary forces (Do et al., 2015; Fu et al., 2014; Grossen et al., 2020; Kyriazis et al., 2020; Robinson et al., 2018) and this may be why some small natural populations can exist without apparent signatures of depression (Robinson et al., 2018). At the individual level, deleterious mutations that are recessive reduce fitness only when they are homozygous (Robinson et al., 2019) as their phenotypic impacts are masked in heterozygotes (Crnokrak & Roff, 1999; García-Dorado, 2012). Thus, the potential genetic load ($Load_P$) harbored by populations due to the number and frequency of deleterious mutations is functionally and mechanistically different from the realized genetic load ($Load_R$) that reduces individual fitness due to homozygosity of deleterious mutations (Mathur & DeWoody, 2021).

Conservation units can be delineated based on genome-wide and functional variation (Funk, McKay, Hohenlohe, & Allendorf, 2012), but here we also consider the potential impact of deleterious mutations (i.e. genetic load). We analyzed individual whole genomes from multiple populations of Montezuma Quail (*Cyrtonyx montezumae;* Fig. 1A), a small Galliform bird in the New World Family Odontophoridae. Like many montane species, Montezuma Quail populations are declining (Harveson, 2009) as their habitat has dramatically reduced in the last century due in large part to land use practices (e.g., domestic livestock grazing) (Brown, 1982; Harveson, 2009; Harveson et al., 2007). In Arizona, Montezuma Quail are abundant and sustainably harvested on public land (Heffelfinger & Olding, 2000) whereas Texas populations are more isolated (Fig. 1A) (23), exist largely on private lands, and have been listed as Vulnerable with no legal hunting for season for decades (Harveson, 2009). We begin by estimating contemporary genome-wide genetic diversity, inbreeding, and the genetic load of putatively deleterious mutations at both the population and the individual level. We do so using samples collected from the large stable Arizona ("potential donor") population as well as small, at-risk ("potential recipient") populations in Texas. Next, we create evolutionary models based on shared genomic regions to age deleterious alleles (i.e. genealogical origin of a mutation) and compare how their frequency changed over time given the demographic history of different Montezuma Quail populations. Finally, we illustrate via simulations how different evolutionary processes such as vicariance and historical demography shape phylogeography, fitness and contemporary load. We then discuss how these inferences may be important for prioritizing donor-recipient pairs for genetic rescue strategies to conserve small threatened populations.

Materials and Methods

Samples & sequencing

Genomic DNA was extracted from 98 Montezuma Quail samples from Arizona (AZ; N=60), West Texas (WTX; N=31), Central Texas (CTX; N=3), and Mexico (MX; N=4) using a modified potassium acetate protocol (Mathur, Tomeček, Heniff, Luna, & DeWoody, 2019; Nicholls, Double, Rowell, & Magrath, 2000). Paired end libraries (2x150 bp) were created using a hyper Library construction kit from Kapa biosystems (Roche) and sequenced on S4 lanes using Illumina NovaSeq 6000. Arizona samples were taken from donated hunter harvested wings initially collected by Randel et al. (2019), whereas lysed blood tissues were taken

3

from West Texas specimens collected from 2018-2020 under a Texas Parks and Wildlife Department Scientific Permit SPR-0410-139 for research. Lysed blood was also taken from Central Texas specimens collected during June and July 2020 under the same permit. Lysed blood was collected from Mexico samples under scientific collection permit number SGPA / DGVS / 06160/19 and were imported to the United States under USDA APHIS permit 140134 issued on 08/December/2019.

Read alignment, preprocessing, and variant discovery

Raw sequences were trimmed to remove adapter sequences and low-quality bases (Phred < 20) using Trimmomatic v.0.36 (Bolger, Lohse, & Usadel, 2014). Filtered reads were mapped to the chromosome level chicken genome (Gallus gallus assembly GRCg6a, genome size = 1.065 Gb) with BWA v.0.7.17 using the *mem* algorithm (Li & Durbin, 2009). We used the Genome Analysis ToolKit (GATK) "Best Practice Workflow" (Auwera et al., 2018) to pre-process our mapped reads. The genomes were sequenced at a mean depth of coverage 11.7x +/- 3.3x per individual (mean +/- SD) (Table S1). We used both GATK v4.1.6 (McKenna et al., 2010) and ANGSD (Korneliussen, Albrechtsen, & Nielsen, 2014) *samtools* model to discover Single Nucleotide Polymorphisms (SNPs) and only retained variants that were concordant between the two methods to avoid biases in variant discovery. For GATK, we used HaplotypeCaller and GenotypeGVCF, and only retained SNPs using hard filters (QD < 2.0 || FS > 60.0 || MQ < 20.0 || MQRankSum <-3.0 || MQRankSum > 3 || ReadPosRankSum < -4.0 || ReadPosRankSum > 4.0 || SOR > 3.0). For ANGSD, we identified SNPs with minQ = 20 and pval = 1e-6. Our initial combined genotype calling and filtering pipeline identified 32,745,572 SNPs in 98 individuals but there was a high variance in the fraction of genotypes missing for AZ individuals, possibly due to degraded DNA and subsequent over-representation of GC-rich sequences (data not shown). To reduce biases in downstream analyses, we removed any individual with [?] 20% missing genotypes. We also only retained variants identified in major chicken autosomes and removed any variants from the sex chromosomes, mitogenome, and unplaced scaffolds. We also removed singletons and private doubletons to avoid biases due to genotyping errors. Ultimately, we analyzed genotypes at 12,943,838 SNPs in 66 individuals (AZ=28, WTX=31, CTX=3, MX=4) placed in chicken autosomes 1-33 (see Table S1). Variant statistics, diversity indices, and heterozygosity (measured as the proportion of heterozygous sites per individual) were calculated using vcftools (Danecek et al., 2011). Chromosome structure is highly conserved among Galliformes with no major rearrangements between the chicken and quail genomes (Morris et al., 2020). This conservation of synteny allowed us to use the chicken recombination map (Groenen et al., 2008) to impute missing genotypes and phase genotypes into haplotypes for Montezuma Quail variants. We used the recombination maps for chicken autosomes 1-28 (except for chromosome 16 due to lack of polymorphic markers; Groenen et al., 2008; Morris et al., 2020) and phased Montezuma Quail SNPs that were mapped to those autosomes using Beagle v.5.1 (Browning, Zhou, & Browning, 2018). Phased haplotypes were used for fineSTRUCTURE analysis and for aging (i.e., estimating the time of origin) deleterious mutations.

Identification of runs of homozygosity (ROHs)

We used PLINK v.1.9 (Chang et al., 2015) to first identify ROHs of size >100 kb using multiple sets of values for parameters associated with scanning window size, SNP density, minimum number of homozygous SNPs within a ROH, minimum number of heterozygous allowed, and maximum internal gap between ROHs. We computed the total proportion of genome within ROH (F_{ROH}) (total genome length = 1,042,168,264 bp) for each set of parameters. To understand the effect of each parameter on ROH identification, we performed a Standardized Regression Coefficients (SRC) analysis to compute the sensitivity indices based on the linear model using *sensitivity* package (Pujol, Iooss, Iooss, & DiceDesign, 2015) in R (R Core Team, 2013). For final identification, we used the parameters –homozyg-window-het 2, –homozyg-snp 50, –homozyg-kb 100, –homozyg-window-snp 20. We used *detectRUNS* package (Biscarini, Cozzi, Gaspa, & Marras, 2018) in R to visualize PLINK output files.

Relatedness, population structure, and shared ancestry

We examined the degree of relatedness among individual pairs by quantifying R0, R1 and KING-robust kinship statistics using IBSRelate (Waples, Albrechtsen, & Moltke, 2019). For estimating population struc-

ture using a model-based approach, we used the genotype likelihood information from all 98 individuals for NGSadmix (Skotte, Korneliussen, & Albrechtsen, 2013) and genotype calls from the best 66 individuals in ADMIXTUREv1.3.0 (Alexander & Lange, 2011). The best value for K was estimated from the rate of change of the likelihood distribution (Evanno, Regnaut, & Goudet, 2005) from the 10 runs in NGSAdmix and 10 cross validation steps in ADMIXTURE.

We next used fineSTRUCTURE (Copenhaver, Lawson, Hellenthal, Myers, & Falush, 2012) to estimate the co-ancestry among individuals using phased haplotypes from 66 individuals (N=132 haplotypes). fineSTRUC-TURE implements the chromosome painting algorithm (ChromoPainter) and reconstructs each chromosome of a "recipient" individual as a patchwork of genetic material inherited from potential donors. Individuals from the same or related populations are expected to share more recent common ancestors than are pairs of individuals from historically separated groups, so co-ancestry would be smaller in the latter case. We considered all individuals as donor-recipient pairs and ran the analysis for each chromosome independently using chicken recombination maps (Groenen et al., 2008). For each chromosome, the fineSTRUCTURE pipeline was executed in 4 steps: the effective population size and mutation rate was first estimated using the expectation maximization (EM) algorithm and then optimized values were used to calculate expected number and the length of inherited segments. Next, an MCMC scheme was used to assign individuals to populations with 10^5 iterations and 50,000 iterations for burn-in, and finally, population trees were built from the converged MCMC runs.

We found of few pair of WTX samples (N=8 pairs) that shared higher co-ancestry than any other pairs (see Results) and indicate potentially familial relationships. Based on R0-R1-KING kinship analysis (Fig. S1), we assigned 1 pair as parent-offspring and 7 pairs as full-siblings (Fig. S2). This was not unexpected given the social structure of quail and the small size of the WTX population (Mathur et al., 2019). To assess any effect of including related individuals in our estimates of genomic diversity, we compared the heterozygosity estimates between Arizona and West Texas populations after removing related individuals from WTX. For each combination of the 24 unrelated WTX samples, we randomly sampled 24 individuals from AZ (with replacement) and compared mean heterozygosity between the two populations using Student's t-test. To test for significance for the observed statistic (t*), we first created a null distribution of the t-statistic by randomly assigning heterozygosity values from the pool of combined values (N=48) into two groups (population 1 and population 2) and estimating the t-statistic by comparing means between the two groups. The distribution was created from 10,000 permutations of random assignments and t-statistic (Fig. S3A). The p-value for observing t* was calculated as p-value = No. of t-statistic values in the null distribution greater than t*/ Total iterations (N=10,000). The p-value for comparison of means between randomly chosen AZ samples and each combination of unrelated WTX individuals ranged from 0-0.0012 (Fig. S3B). Thus, we conclude that even after removing related individuals, the estimate of mean individual heterozygosity was significantly lower for WTX samples as compared to AZ samples.

Functional variant annotation and fitness effect classification

We first used both Ensembl Variant Effect Predictor (VEP) v101.0 (McLaren et al., 2016) and SNPEff v4.3 (Cingolani et al., 2014) to annotate variant sites. Both approaches provide a subjective impact classification of a SNP ("Modifier", "Low", "Moderate", and "High") based on predicted consequences to protein structure and function. We only retained annotations where the impact classes were concordant between the two methods and were without warnings (either due to lack of data or low prediction score). To further classify non-synonymous mutations, we used SIFT (Sorting Intolerant from Tolerant) scores (Ng, 2003) as determined by VEP for all possible amino acid changes in the chicken reference genome (assembly GRCg6a). SIFT scores are generated from sequence homology and the physical properties of amino acids and rank functional (i.e. non-synonymous) variants based on their deleteriousness. We classified the variants into two categories: "Deleterious" and "Weakly deleterious". Deleterious mutations were missense mutations with SIFT score < 0.05, weakly deleterious mutations were missense mutations with SIFT score [?] [0.05, 0.1]. Variants were categorized as synonymous (i.e., neutral) if they had no predicted effect on protein translation or function.

Reconstructing recent demographic history

5

To infer a demographic model that best fits the joint distributions of the observed site frequency spectra (SFS), we simulated SFS under multiple possible demographic models that could best explain the observed SFS. We only considered synonymous mutations for calculating SFS as they are expected to segregate purely due to demographic changes and not due to selective changes. We used Genetic Algorithm for Demographic Model Analysis (GADMA) (Dobrynin, O'Brien, Koepfli, Ulyantsev, & Noskova, 2020) to automatically find the best demographic model for 2 populations: Arizona and Texas. Here, we combined CTX and WTX samples into a single population (TX) due to high co-ancestry between the two populations. We used AZ (N=28) and TX (N=34) samples and projected the observed SFS in [25x25] grid size. To calculate the composite likelihood Akaike information criterion (CLAIC) for goodness-of-fit estimations and confidence intervals (CI) for estimated parameters, we bootstrapped our SFS 100 times. For each model, we started with a series of simple initial structure that includes the number of time intervals that occur before and after each single splitting event and ran forward time simulations using moments within GADMA. The demographic parameters were locally optimized using default algorithms and the global optimization of the best demographic model was done using 20 iterations. We used a generation time of 1 year as Montezuma Quail are effectively an annual species. We first chose the final model based on log likelihood estimates and then performed multiple local optimizations to estimate CLAIC scores and CI.

Estimating genetic load and age of deleterious alleles

Genetic load can be viewed from a gene pool level or at individual level (Fu et al., 2014; Lohmueller, 2014; van Oosterhout, 2020). To distinguish the two perspectives, we use the terms potential load and realized load (see mathematical definitions in Figs. S4, S5, and in Mathur & DeWoody (2021). Potential genetic load ($Load_P$) is defined as the proportion of variants present in the genic sequence that are classified as deleterious or weakly deleterious (Fig. S4). $Load_P$ is similar to segregating load as described by van Oosterhout (2020), but instead of comparing the absolute number of deleterious SNPs, $Load_P$ is conditioned on all non-synonymous, synonymous, and intragenic non-coding SNPs present in a population to avoid biases due to differences in overall genetic diversity and genome size, and to account for all mutations that could possibly affect fitness, directly or indirectly (Mathur & DeWoody, 2021). More diverse populations are expected to have more deleterious mutations (Kim, Huber, & Lohmueller, 2017), but they may or may not have more $Load_P$ depending on whether those mutations significantly outnumber other non-neutral mutations that indirectly affect fitness. To estimate how $Load_P$ is being expressed and decreasing the absolute fitness of individuals, we calculated realized load ($Load_R$) as the proportion of putatively deleterious variants that exist as homozygotes within individual genomes (Fig. S5). An individual would have higher $Load_R$ if it carries higher proportion of deleterious alleles in a homozygous state (e.g., due to consanguineous mating) as compared to another individual where most of the deleterious alleles are heterozygous.

We leveraged the phased data from 132 haplotypes to estimate the age of deleterious alleles by inferring the time to the most recent common ancestor (TMRCA) of deleterious mutation between a pair of haplotypes using Genealogical Estimation of Variant Age (GEVA) (Barton, Albers, & McVean, 2020). GEVA reconstructs genealogical trees of variable sites along each chromosome and estimates the age of any mutation by identifying genomic regions that are identical by descent (IBD) that are shared among haplotypes and broken via recombination. The allele is assumed to be derived from a mutation event in the genome of the common ancestor and shared by all descendent haplotypes. A probabilistic estimate of the time of origin of a mutation is obtained by combining the cumulative distributions for pairs of haplotypes that share the mutation ("concordant pairs") and the pairs that do not ("discordant pairs") (Barton et al., 2020). A mutation is expected to be older than concordant and younger than discordant pairs. We used recombination rate estimates from the chicken (Groenen et al., 2008) and an estimated point mutation rate of 3.14x10⁻⁹ as previously estimate for Montezuma Quail (Mathur & DeWoody, 2021). We used the effective population sizes inferred from EM optimization by ChromoPainter (see above) as a scaling parameter and joint molecular clock to estimate the age of deleterious mutations in different Montezuma Quail populations.

Simulating temporal dynamics of non-neutral mutations

To better understand how mutation load has evolved in Montezuma Quail populations as a result of their

6

recent demography, we performed forward-time simulations using SLiM v.3.6 (Haller, Messer, & Hernandez, 2019) to model the distribution of non-synonymous mutations that arise (via mutation) and segregate conditioned on our best demographic model (see Results). Since SLiM explicitly models each individual within the population, it is memory intensive and thus, we rescaled population sizes to improve simulation performance (Haller et al., 2019). We rescaled our effective populations size (N_e) to 1/10th of that inferred from GADMA, and increased population genetic estimates of mutation rate (μ), recombination rate (r) and selection coefficient (s) 10X so that the products $N_e\mu$, $N_e r$, and $N_e s$ remain unchanged. After rescaling, the demographic model for SLiM ran for 126,700 generations (Fig. S6A). At generation 1, a single ancestral population (p0) was created with size Np0 = 11750 that ran for 117,500 generations (for burn-in) at a constant population size to reach equilibrium. At the end of the 117,500$^{\text{th}}$ generation, p0 expanded instantaneously to 120,000 individuals and evolves at this same size for another 7500 generations until 125,000 generations, where p0 splits in an 80:20 ratio to create two new populations p1 and p2. In our scenario, p1 represents the large Arizona population and p2 represents small Texas population. Immediately after the split, at the end of 125,000$^{\text{th}}$ generation, p2 changes to Np2_1 = 2750 and remains constant for the rest of the simulation whereas p1 linearly declines to Np1_1 = 3600. The migration rates were set to m12 = 4x10^{-3} and m21 = 4x10^{-4} where m12 represents migration from p1 to p2 and vice-versa. To simulate complete isolation between the two populations, we set both migration rates to 0 over the last 200 years, at generation 126,500. Fig. S6A depicts the demographic scenario modelled here.

For our simulations, we recreated the exome of a Galliform-like bird with 33 autosomes that carry a total of 22,412 genes of equal length (1500bp) distributed proportionally to the number of genes found in each chicken chromosome (total genome length simulated = 33,618,000 bp). We assumed no recombination within genes and free recombination between chromosomes using a uniform recombination rate $r = 1x10^{-7}$ crossovers per bp per generation between genes. Mutations arose in our simulation at a rate of $\mu = 3.14x10^{-8}$. We simulated neutral ($s = 0$), deleterious ($s < 0$) and beneficial mutations ($s > 0$) with the ratio of neutral:deleterious:beneficial mutations fixed at 1000:231:2.31. Thus, for every 1000 neutral mutations, 231 mutations were deleterious and only 2.31 mutations were beneficial. The ratio of neutral:deleterious mutations is 1/10th of the corresponding estimate for humans (Kim et al., 2017) and we set the proportion of beneficial mutations to be 1/100th of deleterious mutations. The associated selection coefficients for the deleterious mutations were drawn from a gamma distribution with mean = -0.03 and shape parameter (α) = 0.186 as estimated for humans (Kim et al., 2017) (Fig. S6B). For beneficial mutations, we kept the same distribution and α parameter but changed the mean to 0.03. We did not remove beneficial or deleterious mutations once fixed in the population as they still contribute to absolute fitness. We modeled a complex but biologically relevant dominance scheme that determines the dominance coefficient (h) based on the selection coefficients; the most impactful mutations ($s >= 0.05$) were considered to be completely dominant ($h = 0$) whereas less impactful mutations ($s < 0.05$) have h determined as a step function; neutral mutations were completely additive ($h = 0.5$; Fig. S6C). During the simulations, after the burn-in period, we kept track of genetic diversity indices (heterozygosity, F_{ROH}), mean fitness, and average number of different impact type mutations from random sampling of 1000 individuals from each population (p0, p1, p2) every 50 generations (Kyriazis et al., 2020). The entire simulation post-split (i.e. after 125,000$^{\text{th}}$ generation) was run three times with different initial seed to estimate mean and standard deviation in indices.

Statistical analysis

All pairwise comparisons of means were performed using Wilcoxon rank sum test using false discovery rate (FDR) method of p-value correction. Data in figures are either shown as boxplots or dot plots with explanation in figure legend. Statistical details are provided in Tables S5-S8.

Results

We generated whole genome sequences from Montezuma Quail (*Cyrtonyx montezumae*) populations that vary with respect to contemporary sizes, habitat contiguity (Fig. 1A), and conservation status (Leopold & McCabe, 1957; Harveson, 2009). We studied individual genomes (N=66) from four discrete populations from distinct geographic regions: Arizona (AZ; N=28), West Texas (WTX; N=31), Central Texas (CTX;

N=3), and Mexico (MX; N=4; Fig. 1A). Analysis of 12,943,838 single nucleotide polymorphisms (SNPs) genome-wide revealed the larger Arizona population had significantly higher genomic diversity compared to the smaller and isolated Texas populations (Fig. 1B, Fig. S7). Individuals from the small Texas populations were also more inbred, with significantly more of their genome characterized by runs of homozygosity (F_{ROH}) as compared to the larger Arizona population (Fig. 1C). The presence of many short ROHs (100-200kb; Fig. S8) indicates that contemporary Montezuma Quail carry signatures of historic inbreeding possibly due to founder effects (Ceballos, Joshi, Clark, Ramsay, & Wilson, 2018). More interestingly, we found a genome-wide pattern of reduced heterozygosity (Fig. 1D, Fig. S9), longer ROHs (Fig. 1E) and higher average length of ROHs (Fig. S10) in the small Texas populations. The reduced diversity in Texas quail indicates they have remained small and isolated over evolutionary timescales as seen in Ethiopian and Tibetan wolves (Robinson et al., 2019) or island foxes (Robinson et al., 2018). This is generally in contrast to the recent inbreeding of Isle Royal wolves (Robinson et al., 2019) or Apennine bears (Benazzo et al., 2017) as characterized by bimodal distributions of high diversity regions interrupted by long stretches of ROHs. However, we did find a few Texas individuals with long ROHs (>0.5Mb; Fig. S11) that contain many critical genes (e.g., solute carrier proteins; Table S2). These long ROHs likely reflect recent inbreeding associated with anthropogenic habitat fragmentation and small population size (Harveson, 2009).

Higher potential load in larger populations, higher realized load in smaller populations

Comparing the larger Arizona and the smaller Texas populations, we estimated the contribution of deleterious mutations to the loss of mean population fitness (potential load; $Load_P$) and to the absolute estimated fitness of an individual within the population (realized load; $Load_R$) (Mathur & DeWoody, 2021). $Load_P$ is dependent on number and frequency of deleterious mutations in the population (Fig. S4). Consistent with overall levels of diversity, the larger Arizona gene pool contains significantly more non-synonymous SNPs overall (~1.12X higher; Table S3) and significantly more deleterious mutations than the smaller West Texas gene pool (Fig. S12). Strongly deleterious mutations are more likely to persist in large populations (Kim et al., 2017; van der Valk, de Manuel, Marques-Bonet, & Guschanski, 2019), but we show that these mutations are not over-represented in individual genomes from Arizona as $Load_P$ is not significantly different between the two populations (Fig. 2A). Deleterious mutations arise and exist at significantly lower frequency in the larger Arizona population as compared to the smaller West Texas population (Fig. 2B). The $Load_P$ of weakly deleterious mutations is significantly higher in Arizona than West Texas (Fig. 2A) as weakly deleterious mutations are more likely to survive purging (Grossen et al., 2020) but are still rarer in frequency (Fig. 2B). Thus, larger populations have overall higher $Load_P$ as their deleterious mutations persist at lower frequencies and in heterozygotes, so they evade selection more effectively (Grossen et al., 2020; Robinson et al., 2018). $Load_R$ represents deleterious mutations that are homozygous in individual genomes and characterizes the fitness vulnerabilities and inbreeding depression risks in small populations (Mathur & DeWoody, 2021; Fig. S5). West Texas quail have significantly higher $Load_R$ for both deleterious and weakly deleterious mutations compared to Arizona quail (Fig. 2C).

We found a significant negative relationship between genome-wide heterozygosity and $Load_P$ for deleterious mutations (Fig. S13A). This could mean that larger, more diverse populations are more efficient at purging strongly deleterious mutations than small populations or that larger, more diverse populations are more likely to carry more genic mutations that are neutral or beneficial. Both these processes lead to smaller $Load_P$. However, no such relationship exists for weakly deleterious mutations (Fig. S13B), which indicates that weakly deleterious mutations can persist only when drift overpowers selection. For $Load_R$, we found a significant positive relationship between F_{ROH} and $Load_R$ (Fig. S14). This indicates that despite having a lower proportion of deleterious mutations, inbred populations are more likely to suffer consequences of inbreeding depression as more deleterious mutations are homogenized. The loss of individual fitness due to $Load_R$ can thus accelerate as populations shrink and inbreeding increases leading to longer ROHs (Szpiech et al., 2013).

Montezuma Quail phylogeography and demographic history resembles a ring species

To more fully understand the temporal and spatial distribution of genomic variants associated with load,

phylogeographic and demographic context is necessary. We used phased haplotypes to estimate co-ancestry by comparing regions that are identical by descent (IBD) between pairs of individual Montezuma Quail. Genealogical trees indicate that Arizona and Texas populations are each monophyletic clades whereas the Mexico population is paraphyletic with respect to the Arizona population (Fig. 3A). Furthermore, mean co-ancestry is higher and more homogeneous in Arizona samples compared to Texas (Fig. 3A). The lack of co-ancestry between Arizona and Texas quail buttresses other evidence of historic isolation (Mathur & DeWoody, 2021; Mathur et al., 2019), and the higher co-ancestry between Mexico and Arizona indicates the Arizona population was established after Texas populations were sundered.

We inferred the demographic trajectory of the two divergent Montezuma Quail lineages by simulating the site frequency spectrum of 454,896 synonymous SNPs present in modern Arizona and Texas populations. The most probable demographic model (Fig. 3B, Fig. S15) indicates that the ancestral Montezuma Quail population expanded 87 kya and was relatively large and stable until the Arizona and Texas populations split more recently (16.5 kya; Fig. 3B; Table S4). Texas quail were formed from fewer founders, i.e., they went through a stronger bottleneck, whereas the more diverse Arizona population was formed by more founders (Table S4). The model estimates that Arizona and Texas populations split in an ~ 80:20 ratio and since the split, Texas populations have consistently maintained smaller sizes whereas the Arizona population has been declining to present size. Phylogeographically, our results suggest that Montezuma Quail have evolved much like a ring species complex (Raup & Stanley, 1978). The larger, relatively diverse ancestral population expanded along the Sierra Madre Occidental in the West to form the contemporary Arizona population and ancestors from the Sierra Madre Oriental in the East ultimately formed modern Texas populations (Fig. 3C); the less hospitable habitat of the Central Mexican Plateau likely acted as the major landscape barrier between the two lineages. Our admixture analyses also estimated that all Montezuma Quail individuals most likely were derived from K=2 ancestral populations (S16). Our contemporary samples from Mexico are admixed between the two but are more similar to the western lineage (Fig. 3A, 3D, S16). Geographic conditions suitable for ring species formation with reproductive isolation are rare (Alcaide, Scordato, Price, & Irwin, 2014; Martins, de Aguiar, & Bar-Yam, 2013), but stable terminal populations with historic gene flow can exist if the two populations are locally adapted (Martins et al., 2013). Our model recovered an unequal rate of gene flow between the two populations, indicating a ~10X higher historical migration rate from Arizona to Texas relative to gene flow from Texas to Arizona (Table S4).

Bottlenecks purge ancestral load, but small populations accumulate post-bottleneck load

Do bottlenecks lead to genetic purging? How old are deleterious alleles segregating in contemporary gene pools and how does demography affect their number and frequency? To address these questions, we estimated the temporal origin of deleterious mutations within Arizona and West Texas populations by tracking IBD segments between pairs of phased haplotypes (Barton et al., 2020). We then created linear models to understand the relationship between the age of a mutation and its current frequency within each population. Our results indicate that almost all the deleterious mutations in Montezuma Quail populations originated within the past 150 kya (Fig. S17). Within the last 100kya, more deleterious mutations arose in Arizona (N= 11,102) as than in West Texas (N= 10,308; Fig. 4A). In both populations, there is a reduction in the proportion of deleterious mutations that arose around the time of bottleneck (15-25 kya) but the smaller Texas population continued to more effectively purge deleterious alleles immediately after the bottleneck (10-15 kya) as seen by a drop in mutation density post-bottleneck (Fig. 4B inset). Most of the deleterious mutations are ancestral (i.e. arose pre-bottleneck; Fig. 4B) and thus, are shared between the two populations (Fig. S18). This means genetic purging is most effective immediately post bottleneck as populations are smaller, inbreeding is higher, and founders colonize new environments. Purifying selection against strongly deleterious mutations is a sign of local adaptation and can be mediated by translocations/reintroductions (Grossen et al., 2020). Shared deleterious mutations segregate at similar frequencies in the two populations (Fig. S18) and are likely to rise in frequency at similar rates (Fig. 4C). This means ancestral deleterious mutations that survive natural bottlenecks are more likely nearly neutral thus facing only weak purifying selection (Fig. 4C).

9

These results were confirmed by our exomic forward simulations (Fig. S6) where we detected no loss of fitness after a strong bottleneck due to purging of strongly deleterious mutations in homozygotes via higher inbreeding (Fig. S19). Our simulations show that immediately after the bottleneck, loss of heterozygosity was more pronounced in the large population (Fig. S19A) but inbreeding significantly increased in the small population relative to the large population (Fig. S19B). More interestingly, fitness increased stepwise in both populations, likely due to purging of strongly deleterious alleles (Fig. S19C, D). Beginning ~50 generations post bottleneck, large populations gained heterozygosity via mutation and random mating, but small populations lost heterozygosity due to loss of alleles via drift and selection. Inbreeding stabilized in both populations (Fig. S19B) and fitness generally stabilized in the small population but continued a gradual increase in the large population (Fig. S19C). When looking at the number of mutations of different impacts, we observe that small populations purge strongly deleterious mutations more rapidly (Fig. S19D) than large populations (due to inbreeding depression), but there was no obvious difference in the fate of moderately or weakly deleterious mutations between populations (Fig. S19E, F). We also modeled beneficial mutations ($s > 0$) and observed a similar and gradual accumulation of strongly beneficial mutations in both populations (Fig. S19G). Unlike moderately and weakly deleterious mutations, larger population gained more moderately and weakly beneficial mutations over time (Fig. S19H, I). Our simulations are concordant with our empirical results (Fig. 4) showing that immediately post bottleneck, small populations are more efficient in purging strongly deleterious mutations by means of increased inbreeding that exposes these mutations to selection in homozygotes, ultimately leading to an increase in fitness. Positive selection on strongly beneficial mutations does not play a differential role between small and large populations whereas, both moderately and weakly deleterious and beneficial mutations are higher in larger population, possibly resulting in higher fitness in large population over time.

Most younger mutations that arose *de novo* post-bottleneck (age <10 kya) are privately segregating (Fig. 4B). Arizona has significantly more private deleterious mutations (consistent with more mutations overall; Fig. S18) but we observe a notable increase in the number of deleterious mutations that arose in West Texas population in the last 5,000 years (N= 1250) as compared to Arizona (N= 1099; Fig. 4A, B). Private mutations in the smaller West Texas population always start at higher initial frequencies (1/2N) but unlike shared mutations, have a ~2.5x faster rate of allele frequency increase as compared to the larger Arizona population (Fig. 4C). Not only are these *de novo* deleterious mutations more frequent in the smaller population (Fig. S18), they are also accumulating in individuals as part of $Load_R$ (Fig. S20). Compared to a population that has remained small but does not inbreed due to, for example, mechanisms of inbreeding avoidance or vagaries of habitat (Reed, 2010; Robinson et al., 2018), small populations that have recently inbred can be associated with strongly deleterious mutations (Robinson et al., 2019; Szpiech et al., 2013) and pronounced signatures of inbreeding depression (Benazzo et al., 2017; Crnokrak & Roff, 1999; Robinson et al., 2019), perhaps because of the time lag and environmental stochasticity necessary to effectively purge such deleterious alleles.

Discussion

The preservation of genetic diversity is integral to many modern conservation efforts because of the strong association between genetic diversity and fitness (Bell et al., 2019; Ralls et al., 2020). However, it can be challenging to make strategic decisions that prioritize various conservation scenarios alternatives, such as the best possible source population for genetic rescue. Often, conservation efforts have relied on increasing the overall levels of genetic diversity in the rescued population, mainly by increasing heterozygosity (Frankham et al., 2017). This means translocating individuals from the most diverse source populations (Ralls et al., 2018; Scott, Allison, Field, Averill-Murray, & Shaffer, 2020). However, this "diversity-based regime" (Hansson, Morales, & van Oosterhout, 2021) has been recently challenged based on concerns about the potential for unrecognized genetic load in source populations (Kyriazis et al., 2020; van der Valk et al., 2019; van Oosterhout, 2020).

We think conservation genetics is now at a point where genetic load in various source populations can be evaluated as part of an assessment of overall genetic diversity by using genomic resequencing approaches.

Significant advancements in genome sequencing technology and computational biology now make it feasible to quantify the "deleteriousness" of allelic variants (i.e., their negative fitness effects; (Agrawal & Whitlock, 2012; Cooper et al., 2010)) and then compare the distribution of these effects among populations with different evolutionary histories (Do et al., 2015; Fu et al., 2014). One could argue (correctly) that the detrimental fitness effects of genetic load are usually unknown, but geneticists have long contributed to significant conservation success stories by preserving or enhancing the diversity of genetic variants that had unknown fitness effects (Johnson et al., 2010).

Recent studies of genetic load in small, isolated populations have shown that some species are more robust to being small and isolated (Grossen et al., 2020; Robinson et al., 2018) whereas others may be more likely to succumb to perils of inbreeding depression (Benazzo et al., 2017; Robinson et al., 2019). In this study, we empirically quantified genetic load at both the population and individual level to compare how genetic load manifests itself in in wild populations that vary in size and demographic history (Fig. 2). We also created evolutionary models based on inferred genealogies to estimate the age of deleterious mutations to illustrate how the number and frequency of deleterious mutations varies over evolutionary time (Fig. 4). We demonstrate that genetic load is evolutionarily dynamic because population sizes, the associated efficacy of selection, and the direction and intensity of selection fluctuate with changing environmental conditions (Fig. 3B, Figs. S17-19). We also find evidence that small populations purge deleterious mutations when they undergo strong bottleneck like founding events (Fig. 4A, Figs. S18, 19), as seen in cases of recent colonization (Robinson et al., 2018) as well as reintroductions (Grossen et al., 2020). Our models also reveal that in contrast to large populations, deleterious mutations in small populations are likely to rise in frequency more rapidly (Fig. 4C), when realized load increases due to recent inbreeding (Figs. S14, S20).

Potential load may dictate donor population, but realized load dictates recipient population.

The erosion of genetic diversity can be slowed by assisted gene flow (Frankham, 2015), but the tradeoff between maximizing genetic diversity and minimizing the risk of inadvertently injecting rare deleterious mutations (Kyriazis et al., 2020; Robinson et al., 2018; van der Valk et al., 2019) complicates conservation efforts (Ralls et al., 2020). Translocations based entirely on heterozygosity assessment may be beneficial short term (Scott et al., 2020) but could be harmful in the long run as rare recessive mutations that were first masked as heterozygotes become homogenized in future generations (Hansson et al., 2021). On the contrary, translocating individuals with the fewest deleterious mutations (i.e., prioritizing migrants according to their "load regime") may reduce the probability of load-induced fitness declines, but translocations based entirely on minimizing genetic load would contribute the least genetic variation and may provide few fitness benefits as part of a "rescue" attempt. The best strategy is likely a balanced strategy that considers both the number of deleterious mutations and total estimate of diversity ("diversity + load regime"; (Hansson et al., 2021)). By using $Load_P$ and $Load_R$ (Mathur & DeWoody, 2021) as metrics for comparing genetic load at the population and individual level, respectively, conservationists can more effectively identify potential donor and recipient populations as well as donor individuals for translocation. If individuals from populations with higher $Load_P$ were translocated, they would be more likely to carry deleterious mutations amongst the total influx of new functional variants as compared to individuals from populations with lower $Load_P$. Since $Load_P$ is also inversely proportional to heterozygosity (Fig. S14), populations with lower $Load_P$ are also more likely to contribute neutral or possibly beneficial mutations. We predict that rescue effects will be most pronounced in recently inbred recipient populations where $Load_R$ is high, especially when coupled with assisted gene flow from donor populations with the lowest $Load_P$. In the case of Montezuma Quail, we suggest that if supplementation becomes desirable (e.g., due to increased risks of local extirpations), translocated quail from the Arizona gene pool could reduce $Load_R$ in the recipient population and increase adaptive potential by providing more functional diversity across the genome.

Conclusions

Genetic rescue remains a promising conservation strategy, but we argue that it needs proper evolutionary context before implementation. We reconstructed the demographic history of Montezuma quail populations that vary substantially in contemporary size and show that vicariance associated with a ring distribution has

11

produced the genetic differentiation between extant Texas and Arizona gene pools. Our results show that demographic bottlenecks can lead to genetic purging of ancestral deleterious mutations via purifying selection when the historic population was large. However, once a population becomes isolated and remains small, *de novo* deleterious mutations persist and add to the individual genetic load because purifying selection is ineffective. We conclude that recent inbreeding increases the relative weight of the genetic load in at-risk populations. Our findings provide critical context for evaluating donor populations that can maximize the potential success of rescue efforts like translocations (e.g., by maximizing future adaptive potential while simultaneously minimizing introduction of rare deleterious alleles). We encourage proactive genetic monitoring and argue that evaluations of both overall diversity and genetic load can be used to help prioritize conservation decisions such as the best source populations (or individuals) for genetic rescue. We also think that evolutionary assessments like those conducted herein are beneficial for the holistic conservation of our deteriorating biodiversity.

Data accessibility

The sequence datasets generated during the current study are available in NCBI's Short Read Archive BioProject accession PRJNA703039, BioSample accessions SAMN18007393-18007458 and SRA accessions SRR13748610-13748675. The scripts developed for analysis can be publicly accessed at https://github.com/samarth8392/MQU_EvoGenomics

Authors' contributions

SM, JMT, and JAD conceived and designed the research. LATA and RMP conducted field work and collected key samples. SM performed all wet lab and bioinformatic analyses. SM and JAD led the writing effort and all authors provided input and reviewed the manuscript.

Competing Interest

The authors declare no competing financial interests.

References

Agrawal, A. F., & Whitlock, M. C. (2012). Mutation Load: The Fitness of Individuals in Populations Where Deleterious Alleles Are Abundant.*Annual Review of Ecology, Evolution, and Systematics, 43* (1), 115-135. doi:10.1146/annurev-ecolsys-110411-160257

Alcaide, M., Scordato, E. S. C., Price, T. D., & Irwin, D. E. (2014). Genomic divergence in a ring species complex. *Nature, 511* (7507), 83-85. doi:10.1038/nature13285

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics, 12* (1). doi:10.1186/1471-2105-12-246

Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., . . . DePristo, M. A. (2018). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics, 43* (1). doi:10.1002/0471250953.bi1110s43

Barrett, S. C. H., & Charlesworth, D. (1991). Effects of a change in the level of inbreeding on the genetic load. *Nature, 352* (6335), 522-524. doi:10.1038/352522a0

Barton, N. H., Albers, P. K., & McVean, G. (2020). Dating genomic variants and shared ancestry in population-scale sequencing data. *PLOS Biology, 18* (1). doi:10.1371/journal.pbio.3000586

Bell, D. A., Robinson, Z. L., Funk, W. C., Fitzpatrick, S. W., Allendorf, F. W., Tallmon, D. A., & Whiteley, A. R. (2019). The Exciting Potential and Remaining Uncertainties of Genetic Rescue. *Trends in Ecology & Evolution, 34* (12), 1070-1079. doi:10.1016/j.tree.2019.06.006

Benazzo, A., Trucchi, E., Cahill, J. A., Maisano Delser, P., Mona, S., Fumagalli, M., . . . 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. doi:10.1073/pnas.1707279114

Bijlsma, R., & Loeschcke, V. (2012). Genetic erosion impedes adaptive responses to stressful environments. *Evolutionary Applications, 5* (2), 117-129. doi:10.1111/j.1752-4571.2011.00214.x

Biscarini, F., Cozzi, P., Gaspa, G., & Marras, G. (2018). detectRUNS: detect runs of homozygosity and runs of heterozygosity in diploid genomes.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30* (15), 2114-2120. doi:10.1093/bioinformatics/btu170

Brown, R. L. (1982). Effects of Livestock Grazing on Mearns Quail in Southeastern Arizona. *Journal of Range Management, 35* (6). doi:10.2307/3898250

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. *The American Journal of Human Genetics, 103* (3), 338-348. doi:10.1016/j.ajhg.2018.07.015

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. doi:10.1038/nrg.2017.109

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience, 4* (1). doi:10.1186/s13742-015-0047-8

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2014). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly, 6* (2), 80-92. doi:10.4161/fly.19695

Cooper, G. M., Goode, D. L., Ng, S. B., Sidow, A., Bamshad, M. J., Shendure, J., & Nickerson, D. A. (2010). Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. *Nature Methods, 7* (4), 250-251. doi:10.1038/nmeth0410-250

Copenhaver, G. P., Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of Population Structure using Dense Haplotype Data. *PLoS Genetics, 8* (1). doi:10.1371/journal.pgen.1002453

Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity, 83* (3), 260-270. doi:10.1038/sj.hdy.6885530

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . . Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics, 27* (15), 2156-2158. doi:10.1093/bioinformatics/btr330

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. doi:10.1038/ng.3186

13

Dobrynin, P., O'Brien, S. J., Koepfli, K.-P., Ulyantsev, V., & Noskova, E. (2020). GADMA: Genetic algorithm for inferring demographic history of multiple populations from allele frequency spectrum data. *GigaScience, 9* (3). doi:10.1093/gigascience/giaa005

Ellegren, H., & Sheldon, B. C. (2008). Genetic basis of fitness differences in natural populations. *Nature, 452* (7184), 169-175. doi:10.1038/nature06737

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology, 14* (8), 2611-2620. doi:10.1111/j.1365-294X.2005.02553.x

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics, 34* (1), 487-515. doi:10.1146/annurev.ecolsys.34.011802.132419

Frankham, R. (2005). Genetics and extinction. *Biological Conservation, 126* (2), 131-140. doi:10.1016/j.biocon.2005.05.002

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology, 24* (11), 2610-2618. doi:10.1111/mec.13139

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

Fu, W., Gittelman, Rachel M., Bamshad, Michael J., & Akey, Joshua M. (2014). Characteristics of Neutral and Deleterious Protein-Coding Variation among Individuals and Populations. *The American Journal of Human Genetics, 95* (4), 421-436. doi:10.1016/j.ajhg.2014.09.006

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution, 27* (9), 489-496. doi:10.1016/j.tree.2012.05.012

Garcia-Dorado, A. (2012). Understanding and Predicting the Fitness Decline of Shrunk Populations: Inbreeding, Purging, Mutation, and Standard Selection. *Genetics, 190* (4), 1461-1476. doi:10.1534/genetics.111.135541

Groenen, M. A. M., Wahlberg, P., Foglio, M., Cheng, H. H., Megens, H. J., Crooijmans, R. P. M. A., . . . Andersson, L. (2008). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research, 19* (3), 510-519. doi:10.1101/gr.086538.108

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

Haller, B. C., Messer, P. W., & Hernandez, R. (2019). SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model. *Molecular Biology and Evolution, 36* (3), 632-637. doi:10.1093/molbev/msy228

Hansson, B., Morales, H. E., & van Oosterhout, C. (2021). Comment on "Individual heterozygosity predicts translocation success in threatened desert tortoises". *Science, 372* (6546). doi:10.1126/science.abh1105

Harveson, L. A. (2009). *Management of Montezuma quail in Texas: barriers to establishing a hunting season.* Paper presented at the National Quail Symposium Proceedings.

Harveson, L. A., Allen, T. H., Hernandez, F., Holdermann, D. A., Mueller, J. M., & Whitley, M. S. (2007). Montezuma quail ecology and life history. *Texas quails. Texas A&M University, College Station, USA* , 23-29.

Heffelfinger, J. R., & Olding, R. J. (2000). *Montezuma quail management in Arizona.* Paper presented at the National Quail Symposium Proceedings.

Irwin, D. E., Bensch, S., & Price, T. D. (2001). Speciation in a ring. *Nature, 409* (6818), 333-337. doi:10.1038/35053059

Johnson, W. E., Onorato, D. P., Roelke, M. E., Land, E. D., Cunningham, M., Belden, R. C., . . . O'Brien, S. J. (2010). Genetic Restoration of the Florida Panther. Science, 329(5999), 1641-1645. doi:10.1126/science.1192891

Kim, B. Y., Huber, C. D., & Lohmueller, K. E. (2017). Inference of the Distribution of Selection Coefficients for New Nonsynonymous Mutations Using Large Samples. *Genetics, 206* (1), 345-361. doi:10.1534/genetics.116.197145

Kimura, M., Maruyama, T., & Crow, J. F. (1963). The Mutation Load in Small Populations. *Genetics, 48* , 1303-1312.

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics, 15* (1). doi:10.1186/s12859-014-0356-4

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* . doi:10.1002/evl3.209

Leopold, A. S., & McCabe, R. A. (1957). Natural History of the Montezuma Quail in Mexico. *The Condor, 59* (1), 3-26. doi:10.2307/1364613

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

Lohmueller, K. E. (2014). The distribution of deleterious genetic variation in human populations. *Current Opinion in Genetics & Development, 29* , 139-146. doi:10.1016/j.gde.2014.09.005

Martins, A. B., de Aguiar, M. A. M., & Bar-Yam, Y. (2013). Evolution and stability of ring species. *Proceedings of the National Academy of Sciences, 110* (13), 5080-5084. doi:10.1073/pnas.1217034110

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

Mathur, S., Tomeček, J. M., Heniff, A., Luna, R., & DeWoody, J. A. (2019). Evidence of genetic erosion in a peripheral population of a North American game bird: the Montezuma quail (Cyrtonyx montezumae).*Conservation Genetics, 20* (6), 1369-1381. doi:10.1007/s10592-019-01218-9

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), 1297-1303. doi:10.1101/gr.107524.110

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., . . . Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology, 17* (1). doi:10.1186/s13059-016-0974-4

Morris, K. M., Hindle, M. M., Boitard, S., Burt, D. W., Danner, A. F., Eory, L., . . . Smith, J. (2020). The quail genome: insights into social behaviour, seasonal biology and infectious disease response. *BMC Biology, 18* (1). doi:10.1186/s12915-020-0743-4

Ng, P. C. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research, 31* (13), 3812-3814. doi:10.1093/nar/gkg509

Nicholls, J. A., Double, M. C., Rowell, D. M., & Magrath, R. D. (2000). The evolution of cooperative and pair breeding in thornbills Acanthiza (Pardalotidae). *Journal of Avian Biology, 31* (2), 165-176. doi:10.1034/j.1600-048X.2000.310208.x

Ohta, T. (1992). The Nearly Neutral Theory of Molecular Evolution.*Annual Review of Ecology and Systematics, 23* (1), 263-286. doi:10.1146/annurev.es.23.110192.001403

Pujol, G., Iooss, B., Iooss, M. B., & DiceDesign, S. (2015). Package 'sensitivity'. In: CRAN.

R Core Team. (2013). R: A language and environment for statistical computing.

Ralls, K., Ballou, J. D., Dudash, M. R., Eldridge, M. D. B., Fenster, C. B., Lacy, R. C., . . . Frankham, R. (2018). Call for a Paradigm Shift in the Genetic Management of Fragmented Populations. *Conservation Letters, 11* (2). doi:10.1111/conl.12412

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* . doi:10.1016/j.biocon.2020.108784

Randel, C. J., Johnson, C. Z., Chavarria, P. M., Lopez, R. R., Silvy, N. J., & Tomeček, J. M. (2019). Estimating Montezuma quail hatch date using primary molt at harvest. *Wildlife Society Bulletin, 43* (4), 766-768. doi:10.1002/wsb.1017

Raup, D. M., & Stanley, S. M. (1978). Principles of paleontology (2d ed.). San Francisco: W. H. Freeman.

Reed, D. H. (2010). Albatrosses, eagles and newts, Oh My!: exceptions to the prevailing paradigm concerning genetic diversity and population viability? *Animal Conservation, 13* (5), 448-457. doi:10.1111/j.1469-1795.2010.00353.x

Robinson, J. A., Brown, C., Kim, B. Y., Lohmueller, K. E., & Wayne, R. K. (2018). Purging of Strongly Deleterious Mutations Explains Long-Term Persistence and Absence of Inbreeding Depression in Island Foxes.*Current Biology, 28* (21), 3487-3494.e3484. doi:10.1016/j.cub.2018.08.066

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). doi:10.1126/sciadv.aau0757

Scott, P. A., Allison, L. J., Field, K. J., Averill-Murray, R. C., & Shaffer, H. B. (2020). Individual heterozygosity predicts translocation success in threatened desert tortoises. *Science, 370* (6520), 1086-1089. doi:10.1126/science.abb0421

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating Individual Admixture Proportions from Next Generation Sequencing Data.*Genetics, 195* (3), 693-702. doi:10.1534/genetics.113.154138

Szpiech, Zachary A., Xu, J., Pemberton, Trevor J., Peng, W., Zöllner, S., Rosenberg, Noah A., & Li, Jun Z. (2013). Long Runs of Homozygosity Are Enriched for Deleterious Variation. *The American Journal of Human Genetics, 93* (1), 90-102. doi:10.1016/j.ajhg.2013.05.003

van der Valk, T., de Manuel, M., Marques-Bonet, T., & Guschanski, K. (2019). *Estimates of genetic load in small populations suggest extensive purging of deleterious alleles* . bioRxiv. Retrieved from https://www.biorxiv.org/content/biorxiv/early/2019/07/09/696831.full.pdf

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

Waples, R. K., Albrechtsen, A., & Moltke, I. (2019). Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. *Molecular Ecology, 28* (1), 35-48. doi:10.1111/mec.14954

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in Ecology & Evolution, 30* (1), 42-49. doi:10.1016/j.tree.2014.10.009

Figure 1. Genomic diversity and inbreeding in Montezuma Quail.(A) The Montezuma Quail (inset) is a small gamebird predominantly distributed in Mexico with peripheral populations in the U.S. We analyzed 66 whole genomes from Arizona (AZ=28, pink circle), West Texas (WTX=31, purple arrow), Central Texas (CTX=3, orange arrow), and Mexico (MX=4, blue circle). (B) Boxplots showing mean genomic heterozygosity estimates indicate the isolated populations in Texas are significantly less heterozygous than the AZ population. (C) Small, isolated Texas populations have significantly more of their genome in runs of homozygosity (ROH), with CTX individuals having highest inbreeding coefficient (F_{ROH}). (D) Sliding

window analysis of heterozygosity across non-overlapping 1kb windows corresponding to chicken chromosome 1 also reveal an overall reduction in genomic heterozygosity in the WTX population (break represents centromere; see Fig. S2 for all chromosomes). (E) The distribution of ROH length across individual genomes (each column along the x-axis). Individuals in WTX and CTX populations have longer ROHs, on average, compared to AZ or MX individuals. Range map was acquired from the IUCN database. In boxplots, center line indicates median; box edges are upper and lower quartiles; whiskers represent 1.5x interquartile range and solid black points represent outliers. (* p<0.05, ** p<0.01, *** p<0.001). Montezuma Quail image courtesy: Bob Gress, BirdsInFocus.

Figure 2. Potential load and realized load in Montezuma Quail.(A) Potential load, the proportion of all genic mutations that are deleterious, is similar overall between the larger Arizona and smaller West Texas population, but Arizona has a greater potential load of unpurged weakly deleterious mutations. In boxplots, center line indicates median; box edges are upper and lower quartiles; whiskers represent 1.5x interquartile range and solid black points represent outliers. (B) Mean (circle) and SE (error bar) minor allele frequency (MAF) of deleterious mutations in two populations. As expected, MAF was significantly higher in small Texas population for both deleterious and weakly deleterious mutations. (C) Realized load, the proportion of deleterious mutations that exist as homozygotes in individual genomes, is higher in the small West Texas population and may contribute to inbreeding depression. (* p<0.05, ** p<0.01, *** p<0.001)

Figure 3. Phylogeography and demographic history of Montezuma Quail. (A) Genealogical tree (left) and co-ancestry heat map for sampled populations (right). The three major clades are shown within black boxes and the Central Texas samples are shown in the dotted box. Both Arizona and Mexico samples share higher within population co-ancestry; Texas quail are more heterogeneous. Lack of co-ancestry between Texas and Arizona populations (bottom right diagonal) indicates long term isolation. (B) Our demographic model of Arizona and Texas populations indicates the ancestral population (NA0) underwent demographic expansion (NA1) around 90 kya followed by a split 17 kya. Texas populations have maintained small sizes (NTX) since their split whereas contemporary Arizona populations (NAZ) formed after a series of bottlenecks. We assumed a generation time of 1 year for Montezuma quail; arrow size is proportional to migration rate between AZ and TX. (C) The phylogeography of Montezuma quail resembles that of a ring species, with an ancestral population expanding along the Sierra Madre Occidental in the west to form the contemporary Arizona population and a smaller eastern expansion along the Sierra Madre Oriental forming the modern Texas populations (left). A schematic showing the Central Mexican Plateau as the major landscape barrier between the two lineages is also shown (right). (D) Admixture analysis reveal subdivisions within Arizona, coancestry between AZ and MX, and pronounced differentiation between AZ and TX populations. K= no. of ancestral populations.

Figure 4. Temporal dynamics of genetic load. (A) The age distribution of contemporary deleterious mutations in Arizona and West Texas over the last 100kya. Much of the contemporary load (~50%) originated >50 kya in the ancestral population, whereas ~30% of the deleterious alleles arose in the last 10-25 kya since the demographic split (indicated by the dotted line). Despite its much smaller size, the West Texas population contains more younger (<5 kya) deleterious mutations that have arisen de novo (red bar). (B) The age distributions of deleterious mutations that are shared between Arizona and WTX or are privately segregating over the past 150 kya and 25 kya (inset). The dip in the density of deleterious mutations around ~17kya (black bar in inset) suggests historic purging in West Texas due to a founding bottleneck (Fig. 2A) but purging of the recent deleterious mutations (age < 5kya) has been less efficient. (C) Linear model of mutation age and frequency predict similar trajectories for shared deleterious mutations (upper panel) but private deleterious mutations (lower panel) in West Texas have faster rate of frequency increase. The higher intercept for the West Texas population is expected given its small size as new mutations always have an initial frequency of 1/2N. These results demonstrate that small populations ineffectively purge de novo deleterious mutations.

Hosted file

`Figure1.pdf` available at https://authorea.com/users/311018/articles/528252-an-evolutionary-