Genomic health is dependent on population demographic history

Eric Wootton[1], Aaron B.A. Shafer[2]

[1]Biochemistry & Molecular Biology, Trent University, K9J 7B8 Peterborough, Ontario, Canada

[2]Environmental and Life Sciences Graduate Program, Trent University, 7K9J 7B8 Peterborough, Ontario, Canada

Corresponding authors:

Eric Wootton: ericwootton@trentu.ca

Aaron Shafer: aaronshafer@trentu.ca

Abstract

Current genetic methods of population assessment in conservation biology have been challenged with genome-scale data, particularly by the quantitatively novel insights provided by genome-wide analyses. These include assessments of runs-of-homozygosity (ROH), genomic evolutionary rate profiling (GERP), and mutational load, all of which are direct measures of genomic health. Here, we aim to elucidate the relationships between these measures and assess the genomic health of three divergent ungulates: the white-tailed deer, caribou, and mountain goat. The white-tailed deer is currently expanding following near extirpation, while caribou are in the midst of a significant decline. Mountain goats remain stable after a large historical bottleneck. We assessed genome-wide signatures of inbreeding using the inbreeding coefficient F and % ROH (F_{ROH}) and identified evolutionarily constrained regions with GERP. Mutational load was estimated by identifying mutations in highly constrained elements (CEs) and by sorting intolerant from tolerant (SIFT) mutations. Our results show that F and F_{ROH} are significantly higher in mountain goats than in caribou and white-tailed deer. Given the extended bottleneck and low N_e of the mountain goat, this supports the idea that the genome-wide effects of demographic change take time to accrue. Similarly, we found that mountain goats possess more highly constrained CEs which are indicative of greater purifying selection. This is further reflected by fewer mutations in CEs and deleterious mutations identified by SIFT. In contrast, white-tailed deer presented the highest mutational load with both metrics. Our results demonstrate that extended bottlenecks may lead to reduced diversity and increased F_{ROH} in ungulates, but not necessarily the accumulation of deleterious alleles. This may be due to the purging of deleterious alleles in small populations. This study empirically demonstrates the

relationships between different measures of genomic health in ungulates and highlights the need to consider multiple genomic health metrics during conservation assessments.

Introduction

Internationally, organizations like the International Union for Conservation of Nature (IUCN) rely primarily on population sizes and trends when assigning wildlife conservation statuses (IUCN Standards and Petitions Committee, 2019). The predictive relationship between population size (N_C) and genetic diversity is challenging, as N_C, for example, tends to be substantially more variable than diversity metrics like nucleotide diversity (π), referred to as Lewontin's paradox (Ellegren & Galtier, 2016; Lewontin, 1974), and therefore rarely accounts for recent demographic history and the genetic health of populations. While genetic diversity estimates generally correlate with IUCN status (Peart et al., 2020; Petit-Marty et al., 2021), the relationship (i.e., direction) is varied and predictive association is low (Schmidt et al., 2022). The unprecedented loss of global biodiversity (Pascual et al., 2021) and ambiguous associations (Schmidt et al., 2022) necessitates the need for conservation-informative genetic metrics.

Both novel genomic metrics and functional versus presumed neutral variation have garnered attention in the conservation genetics community. The importance of neutral genetic diversity for example, despite a long history (e.g., DeWoody et al., 2021; Kohn et al., 2006), was recently called into question (Teixeira & Huber, 2021). Citing Lewontin (1974), Teixera & Huber (2021) argued for a renewed focus on functional variation and gene-environment interactions, which prompted two high-profile responses (Garcia-Dorado & Caballero, 2021; Kardos et al., 2021). Kardos et al. (2021) in particular echoed earlier concerns of [functional] gene-targeted conservation (i.e. Kardos & Shafer, 2018), and highlighted some potential problems in Teixeira & Huber's simulations. Kardos et al. (2021) strongly argued against

abandoning genome-wide neutral markers. Supporting this view would be empirical work from Peart et al. (2020) that showed Tajima's *D*, a composite statistic derived from Watterson's θ and π (Tajima, 1989), appears predictive of conservation status, further reiterating the value of classic neutral summary statistics derived from genome-wide sequence.

The high correlation between nucleotide diversity and microsatellite heterozygosity means genome-scale data reflect smaller targeted panels of loci (Väli et al., 2008). This fact is ultimately what formed the premise of Shafer et al. (2015) that argued for the use of quantitatively novel genomic metrics when it comes to conservation. Runs-of-homozygosity (ROH), genomic evolutionary rate profiling (GERP), and mutational load are metrics dependent on genome-scale data and have only recently been applied to species and populations of concern (Grossen et al., 2020; van der Valk et al., 2019; von Seth et al., 2021). ROHs are long stretches of homozygosity that arise as a result of extensive inbreeding (Ceballos et al., 2018), and both reflect the true inbreeding coefficient and can identify genes underlying inbreeding depression (Kardos et al., 2016). GERP profiling can be used to identify evolutionarily constrained genomic elements, defined as regions with fewer substitutions than expected under neutral evolution (Davydov et al., 2010), which are often the subject of selective pressures (Cooper et al., 2005; Davydov et al., 2010). Any mutations that appear at highly conserved sites are generally assumed to be deleterious (Cooper et al., 2005; Davydov et al., 2010). Likewise, tolerant and intolerant mutations – also referred to as SIFT scores (Vaser et al., 2015) – that are considered of human clinical interest (P. Kumar et al., 2009) – have now been applied to natural populations (Robinson et al., 2016). ROH, GERP and SIFT scores collectively allow for inferring the mutational load and genomic health of a population or species.

While quantitatively useful, the relationships between these metrics and conservation status and diversity metrics are generally not well defined and at-times counter to expectation. van der Valk et al. (2019) showed a weak relationship between conservation status and mutational load across mammalian taxa; the highest mutational loads were in fact seen in the most abundant and diverse species. Similarly, species with the highest inbreeding (ROH) also had the lowest mutational load and GERP scores (van der Valk et al., 2019). This pattern could be due the purging of deleterious alleles in small populations, which is consistent with simulations by Kyriazis et al. (2021). van der Valk et al. (2019) argued that this process has implications for genetic rescue and suggested that individuals be selected for translocation based on low mutational load and not diversity. Work on Sumatran rhinoceros added another layer of nuance, with populations showing low to moderate levels of inbreeding (ROH), but a relatively high mutational load (von Seth et al., 2021). In contrast, the Alpine ibex, that has experienced successive demographic bottlenecks that led to reduced diversity and increased ROH, but not the accumulation of highly deleterious alleles (Grossen et al., 2020). Thus, there seems to be a link between historical demography, ROH, and mutational load. Understanding such processes has important implications for both predicting long-term viability, but also selecting animals for genetic rescue.

Here, we analysed signatures of genomic health across populations of three divergent ungulates to elucidate these relationships. The caribou (*Rangifer tarandus*) is currently experiencing a significant population declines (N_C) prompting international concern (COSEWIC, 2015). In Quebec, the five major herds have seen drastic declines from their highest estimates ranging between 10% and 75%, depending on the herd (COSEWIC, 2015). The white-tailed deer (*Odocoileus virginianus*), in contrast, is expanding northward and has accordingly high genomic

diversity (Kessler et al., 2022); while white-tailed deer currently number in the 10s of millions (Kauffman et al., 2018), they were historically extirpated across much their range due to colonial harvest and forestry practices (Kauffman et al., 2018); meaning this demographic expansion is very recent. The mountain goat (*Oreamnos americanus*) underwent a massive bottleneck during the last glacial maximum (Martchenko et al., 2020) but has since remained stable with a range-wide population of ~100,000 individuals (Festa-Bianchet & Côté, 2007). These three species, contrasting in conservation status, historical and demographic trajectories, present useful comparisons to explore patterns of novel genomic metrics as they pertain to conservation status. In line with recent empirical work, we predicted that populations with low diversity and high inbreeding do not necessarily have higher mutational load and that genome-wide effects of demographic change like ROH take time to accrue. Similarly, we expected to see greater purging of deleterious alleles in populations with long-term small N_e. Collectively, this means that poor genomic metrics of health would not correlate with current negative population trends.

Methods

Caribou samples were obtained from 67 individuals across Quebec and Labrador (Dedato et al. 2022); 53 white-tailed deer were sampled across their North American range (Kessler et al. 2022). Additionally, 35 mountain goats were sampled across their native range from Alaska to as far south as Idaho (Martchenko et al. 2022). See Figure 1 for an illustration of sampling distribution. All samples were sequenced with Illumina HiSeqX technology to approximately 5X coverage. Non-target sequences were trimmed from raw FASTQ files with Trimmomatic v.0.36 (Bolger et al., 2014), and reads were mapped to their reference genome using BWA-MEM v.0.7.17 and SAMtools v.1.15 (Li et al., 2009; Li & Durbin, 2009). Duplicate mapped reads were marked and removed using Picard v.2.23.0 (Broad Institute, 2019). Uniquely mapped reads were

subset by Sambamba v.0.8.2 (Streit et al., 2012), and indels were realigned with GATK v.3.8 (Van der Auwera et al., 2013). We estimated genotype likelihoods using ANGSD v.0.934 using the GATK model: -doVcf 1 -doGeno -4 -doPost 1 -gl 2 -SNP_pval 1e-6 -minMapQ 20 -minQ 20 -doCounts 1 -skipTriallelic 1 -doGlf 2 (Korneliussen et al., 2014). Hard calls required for analyses were performed by Beagle v.5.1 (Browning et al., 2018) with the following parameters: impute=false overlap=3000 window=100000 gprobs=TRUE parameters. Posterior genotype probabilities were estimated using allele frequencies, and only genotypes with a posterior probability above 0.95 were called. Tajima's *D* and nucleotide diversity (π) were calculated using VCFtools v.0.1.13 with windows of 10 Kb and using all positions, regardless of allele frequency (Danecek et al., 2011).

Inbreeding and Homozygosity

A pruned to a subset of SNPs in approximate linkage equilibrium which was obtained using PLINK v.1.9 with a window size of 50 SNPs, a window shift size of 10 SNPs, and a pairwise (R^2) linkage disequilibrium threshold of 0.1 (Chang et al., 2015; Purcell et al., 2007). We then identified runs of homozygosity (ROH) on scaffolds greater than 10 Mb in length and with a minor allele frequency cut-off of 0.05. Sliding windows of 50 SNPs were used, and each ROH was defined as a sequence at least 100 Kb long with at least 50 homozygous SNPs. Each ROH required a minimum density of 1 SNP per 50 Kb and no more than 5 missing or heterozygous calls in a window. Any SNPs more than 1 Mb apart within a segment were used to split that segment in two. Once identified, each ROH was classified by size, and the percent ROH (F_{ROH}) for each scaffold and individual were calculated. PLINK parameters were taken from Foote et al. (2021) with output validated using the program ROHan (Renaud et al., 2019). ROHan was run on the same data with the same ROH size definition of 100 Kb and a transition to transversion

ratio (Ts/Tv) of 2.25. This ratio was calculated using BCFtools v.1.14 and the genotypes provided by ANGSD (Li et al., 2009). The inbreeding coefficient (*F*) for each individual was also estimated using PLINK and the –het flag.

Evolutionary Constraint and Mutational Load

Twenty ungulate reference genomes from the NCBI GenBank genome database (Table S1) were selected to represent most extant families with genomes available in both the orders of Perissodactyla (odd-toed ungulates) and Artiodactyla (even-toed ungulates). Of our focal species Caribou did not have an available annotation; here we used a mixture of *de novo* and homology-based predictions. We aligned protein sequences from *Bos taurus, Equus caballus, Homo sapiens, and Ovis aries* to the caribou genome using blastx v.2.7.1 (Altschul et al., 1990). These sequences were all obtained from Ensembl release 89 (Hunt et al., 2018). We also aligned transcripts from the NCBI mammalian RefSeq transcipt database v.1.0 using BLAT v.1.0 to help identify exons. Caribou RNAseq data (SRX12486288) from the NCBI Sequence Read Archive was aligned to the genome with HISAT v.2.10 for *de novo* prediction (Kim et al., 2015). Using these data, genes were predicted by Augustus v.3.1.1 as well as GeneMarkES v.4.69 (Lomsadze et al., 2005; Stanke et al., 2006). We generated a set of consensus genes using EVidenceModeler v.1.1.1 with the following weighting scheme (Haas et al., 2008): transcript alignment via BLAT (10x weight); protein alignment via blastx (5x weight); gene prediction via Augustus (1x weight); gene prediction via GeneMarkES (1x weight). PASA v.2.3.3 was used last to refine our gene identifications (Haas et al., 2008).

Constrained regions for each individual were identified using the Genomic Evolutionary Rate Profiling (GERP++) ; this approach requires a phylogenetic tree which we constructed for all 23 species using TimeTree (Figure S1) (S. Kumar et al., 2017). The individual fastq files of all individuals of our three species were aligned to each of the 20 references in 50 bp segments using BWA-MEM. All duplicate alignments were marked and removed with Picard. Then, the alignments of each individual were concatenated into one multiple alignment file (MAF), resulting in 155 total MAFs. The gerpcol function included in GERP++ was run on each MAF to generate rejected substitution (RS) scores for each alignment column. RS scores are used as a metric for constraint intensity and are calculated by subtracting the estimated evolutionary rate at each position from the neutral rate (Cooper et al., 2005; Davydov et al., 2010). Both rates were estimated by gerpcol using the aforementioned phylogenetic tree (Figure 2). RS scores quantify the number of differences between the expected and observed number of SNPs at a site, and a positive score is indicative of constraint (Cooper et al., 2005; Davydov et al., 2010). An RS score greater than 2 represents a highly constrained region whereas a score lower than 1 is putatively neutral and discarded (Cooper et al., 2005; Davydov et al., 2010). gerpcol was run with the previously calculated Ts/Tv ratio of 2.25.

We also used GERP++ to identify constrained elements with the gerpelem function. To estimate mutational load, we determined the number of SNPs within these constrained elements using BEDtools v.2.30.0 (Quinlan & Hall, 2010). These SNPs are putatively deleterious due to the common functional importance of constrained regions (Cooper et al., 2005; Davydov et al., 2010). For each species, the functional consequence of each SNP was determined using Ensembl's Variant Effect Predictor (VEP) (McLaren et al., 2016). The deleteriousness of each SNP across the genome was then estimated using functional tolerance scores provided by SIFT

and SIFT4G (Vaser et al., 2015). Using multiple alignment information, SIFT and SIFT4G calculate normalised probabilities where substitutions with scores lower than 0.05 are predicted to be deleterious to protein function (Vaser et al., 2015).

For all individual metrics above we quantified the correlation to sample location (latitude, longitude). Because white-tailed deer were extirpated across much of their range, we used translocation data from McDonald & Miller, (2004) to determine if these metrics were correlated to restocking efforts in the United States. We examined correlations with stock start and end dates (year), estimated population low, number of individuals stocked, and the number of states involved in stocking.

Results

Diversity and Inbreeding

Nucleotide diversity (π) was found to be highest in deer (mean = 0.00771) and lowest in mountain goats (mean = 0.00101) (Figure 1). Tajima's *D* was highest in mountain goats (mean = 0.625) and lowest white-tailed deer (mean = -0.645) (Figure 1). Mountain goats were found to have significantly higher F_{ROH} (mean = 0.038) than both caribou (none detected) and white-tailed deer (mean = 0.0020) (Figure 2a). This pattern is reflected in the relative values of *F* where mountain goats had the highest value (mean = -0.018) in comparison to caribou (mean = -0.44) and deer (mean = -0.27) (Figure 2b). For the mountain goat, most ROHs were found to be in the 500-1000 Kb range whereas deer possessed ROHs almost exclusively below 300 Kb (Figure 3). Despite a much higher number of ROHs overall, mountain goats possessed significantly fewer in the 100-300 Kb range than deer (Figure 3). Both measures of inbreeding were found to be

correlated in mountain goats (Pearson's $r = 0.840$, $p < 0.05$) and deer (Pearson's $r = 0.395$, $p < 0.05$) (Figure S5), noting again there were no in F_{ROH} caribou preventing a similar assessment.

Evolutionary Constraint and Mutational Load

Mountain goats had the highest number of constrained elements (CEs) (mean = 2926) while caribou have the fewest (mean = 935; Figure S3). Highly constrained elements followed a nearly identical distribution between species with all three having approximately 50% as many CEs with RS > 2 (Figure S3). In addition to having the most CEs, mountain goats showed slightly longer CEs on average (Figure S4). The three species showed little variability in average CE RS-score (Figure S2, S4). Despite having significantly fewer CEs, white-tailed deer had a larger number of SNPs in CEs (mean = 411) followed by mountain goats (mean = 212) and caribou (mean = 74) (Table 1). The relative differences between species remained the same for highly constrained CEs (Table 1). We observed that a much higher fraction of constrained positions were deleterious in deer (mean = 0.625%) come to caribou (mean = 0.326%) and mountain goats (mean = 0.121%; Figure 4).

We found 42.1% of coding sequence consequences identified by VEP to be missense variants in white-tailed deer, 49.3% in caribou, and 57.4% in mountain goats (Figure 5a). Despite this distribution, white-tailed deer had the highest number of intolerant mutations (mean = 6852) identified by SIFT followed by caribou (mean = 4944) and mountain goats (mean = 3190) (Table 1). Similarly, we observed white-tailed deer to have the highest fraction of genes with deleterious mutations (mean = 33.0%) while caribou (mean = 14.9%) and mountain goats (mean = 14.3%) had the lowest (Figure 5b). None of these metrics were correlated to the number of genes in the annotation (Table 1).

Lastly, we found F and latitude to be moderately correlated in caribou (Pearson's r = -0.466, $p < 0.05$) and deer (Pearson's r = -0.513, $p < 0.05$; Figure S6), but not in mountain goats. There were no significant correlations between our individual genome health metrics and restocking data in white-tailed deer (Pearson's $r < 0.2$, $p < 0.05$). No significant differences could be found between deer from states with translocations and those from states that have had no translocations.

Discussion

The interplay between nucleotide diversity (π), inbreeding, mutational load, and census size (N_C), is nuanced and full of contrasting examples. Here, we demonstrate that mutational load may be lower than expected in populations with high inbreeding and low diversity. Our observed patterns suggest that the purging of deleterious alleles in these cases is stronger in small populations (Huber et al., 2020), and therefore would reduce risks from inbreeding depression. This supports our hypotheses that low diversity and high inbreeding do not on their own necessitate special conservation status, and that purging is occurring in populations with low N_e. We also predicted that genomic signatures of demographic events take time to accrue, which is supported by evidence of F_{ROH} and CEs requiring extended bottlenecks to surface and increase. Caribou have seen dramatic recent decreases in N_C yet have no detectable ROH; in contrast, mountain goats show elevated ROH likely due to the ancestral nature of their bottleneck (Martchenko et al., 2020). The varied relationships between different measures of genomic health illustrates and lack of correlation to the population trends that are typically used for status assignment highlights the value of this comparative framework (see also Huber et al., 2020). This information is especially valuable for small populations of conservation concern, where

mutational load combined with ROH and CE assessments is probably the most informative way forward, in addition to metrics like Tajima's *D* (Peart et al., 2020).

Inbreeding Effects

Genomic consequences stemming from demographic take time to accumulate (e.g. Grossen et al., 2020; von Seth et al., 2021). Caribou are currently in the midst of a massive demographic bottleneck resulting in an extremely low contemporary effective population size (N_e) and nucleotide diversity (Dedato et al., 2022). In contrast, mountain goats underwent a large historical bottleneck during the last glacial maximum, and while they have not recovered, they have remained relatively stable, resulting in a moderately low N_e (Martchenko & Shafer, 2022). Low diversity stemming from bottleneck is supported by positive values of Tajima's *D* in these two species, which indicate a lack of rare alleles and population contraction (Tajima, 1989). The lower π in mountain goats is clearly due to their extended bottleneck that would have exacerbated the effects of inbreeding, but also their historically lower census sizes. These effects are reflected by F_{ROH}, where mountain goats were found to have 3.8% of the genome in ROH whereas caribou had virtually no detectable ROH segments. Here, despite a less drastic decline in N_C, mountain goats have accumulated a significantly greater F_{ROH} than caribou, and in comparison to white-tailed deer, mountain goats have developed significantly longer and more numerous ROHs, which further outline the effects of a prolonged bottleneck. Caribou simply have not been subject to a low N_C for long enough to approach mountain goats or even differentiate themselves from deer in their levels of inbreeding. Of note, F_{ROH} was more correlated to F after an extended bottleneck (i.e., mountain goats vs. white-tailed deer; Figure S5), which further implies that ROHs require both time and demographic change (Gillespie, 2004) to accumulate. Overall, these results support the idea that genome-wide effects of

demographic change will not accrue unless given enough time; F and F_{ROH} are therefore unlikely to align with IUCN conservation decisions, similarly noted by Schmidt et al. (2022).

Mutational Load and Purging

The relationship between N_e, inbreeding, π, and mutational load are variable (Grossen et al., 2020; van der Valk et al., 2019; von Seth et al., 2021). Empirical studies on mammals have led us to predict that low diversity and high inbreeding do not necessarily result in the accumulation of deleterious alleles (Grossen et al., 2020; Robinson et al., 2018, 2018; van der Valk et al., 2019). Here, we found a much greater number of putatively deleterious mutations in white-tailed deer than in mountain goats and caribou by identifying CE-bound SNPs with GERP. These results were confirmed by the deleterious alleles identified by SIFT that showed a similar gap between white-tailed deer and the other species, regardless of the annotation quality (Table 1). This means that despite lower diversity and small N_e, caribou and mountain goats have relatively low mutational loads; further, despite a relatively low number of CEs, deer still possessed significantly more CE-bound SNPs indicative of deleterious mutations. Mountain goats, in contrast, were found to have the highest amounts of inbreeding but still not accumulated many deleterious alleles. This reiterates the finding that inbreeding does not necessarily result in high mutational load; in the case of Caribou, if N_C can return (e.g., Peart et al. 2020) or if purging can occur (e.g., Huber et al., 2020), the genetic cost to the current crash in N_C could be minimized. This, however, remains an open question.

The purging of deleterious alleles in small populations is rarely observed (Hedrick & Garcia-Dorado, 2016; Robinson et al., 2018; Xue et al., 2015), but there is evidence that it may occur in the presence of severe bottlenecks (Grossen et al., 2020). Mountain goats were shown to have low mutational load in conjunction with generally low diversity, while caribou present

intermediate profiles. This would be consistent with N_C numbers pre-20th century for caribou, that were presumably much higher (recognizing still that natural cycles were occurring). Reductions in N_e from severe bottlenecks can increase the fitness effects of recessive deleterious alleles by increasing homozygosity (Garcia-Dorado, 2012, 2015); this strengthens the pressure of negative selection and can ultimately lower the mutational load of a population (Garcia-Dorado, 2012, 2015; Glémin, 2003).

Purifying selection should remove SNPs in functionally important regions and constrain these areas against change, which is why we observe less SNPs than expected in CEs (Cooper et al., 2005; Davydov et al., 2010). The mountain goat's longer and more numerous CEs relative to the other species are possibly indicative of stronger purifying selection, the driving force behind constraint. Caribou in contrast, are now in the midst of a bottleneck and have a smaller N_e than deer; thus purging has either yet to take place or selection is less efficient, or a combination of both. The trajectory of caribou, were they to not go extinct, would be predicted to be towards the patterns of mountain goat based on mammalian wide patterns and profiles (Cooper et al., 2005; Huber et al., 2020). It is also important to note that purifying selection reduces nucleotide diversity by eliminating sites under direct selection and those linked through background selection (Charlesworth, 1996; Cvijović et al., 2018), which partially explains patterns of diversity seen here (Figure 1; Buffalo, 2021).

This work has some implications for conservation decisions, and more broadly the expected genomic consequences of numerical population declines. Sumatran rhinos for example have a very similar historical N_e to that of mountain goats (von Seth et al., 2021), but considerably more ROH. While purging might have played some role in some populations (von Seth et al., 2021), the overall negative effect of autozygosity and [likely] continued presence of

high frequency deleterious alleles (e.g., Khan et al., 2021) was probably insurmountable from a population viability standpoint. Thus, we posit that high mutational load must be accompanied by high ROH to be of conservation concern, similar to what was observed in woolly mammoth (Rogers & Slatkin, 2017). Further, genetic rescue decisions should not rely solely on inbreeding and diversity metrics, especially when their relationships to fitness are not so straightforward (Schmidt et al., 2022). This is also true for assessing restocking efforts, as we demonstrate that translocations are not necessarily related to population health metrics. Collectively, this provides further evidence of the time needed for genome-scale changes to accrue after significant changes in demography and how a holistic view of genomic metrics is required for an accurate assessment of genome health.

Acknowledgements

We wish to acknowledge that the work for this study was carried out on the traditional territory of the Mississauga (Michi Saagiig) Anishinaabeg people. This work was supported by a Natural Sciences and Engineering Research Council Discovery Grant, Canada Foundation for Innovation: John R. Evans Leaders Fund, and Compute Canada Awards to ABAS; Eric Wootton was supported by three NSERC USRAs. We thank Daria Martchenko, Morgan Dedato and Camille Kessler for help with data acquisition. Sample providers are listed in Dedato et al. (2022), Martchenko & Shafer (2022), and Kessler et al. (2022).

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

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

Broad Institute. (2019). *Picard Toolkit*.

Buffalo, V. (2021). Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's Paradox. *ELife*, *10*, e67509. https://doi.org/10.7554/eLife.67509

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 2018 19:4*, *19*(4), 220–234. https://doi.org/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), 7. https://doi.org/10.1186/S13742-015-0047-8

Charlesworth, B. (1996). Background selection and patterns of genetic diversity in Drosophila melanogaster. *Genetical Research*, *68*(2), 131–149. https://doi.org/10.1017/s0016672300034029

Cooper, G. M., Stone, E. A., Asimenos, G., Comparative, N., Program, S., Green, E. D., Batzoglou, S., & Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. *Genome Research*, *15*(7), 901–913. https://doi.org/10.1101/gr.3577405

COSEWIC. (2015). *COSEWIC Assessment and Status Report on the Caribou Rangifer tarandus, Newfoundland population, Atlantic-Gaspésie population, Boreal population in Canada*. https://species-registry.canada.ca/index-en.html#/documents/2769

Cvijović, I., Good, B. H., & Desai, M. M. (2018). The Effect of Strong Purifying Selection on Genetic Diversity. *Genetics*, *209*(4), 1235–1278. https://doi.org/10.1534/genetics.118.301058

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Davydov, E. V., Goode, D. L., Sirota, M., Cooper, G. M., Sidow, A., & Batzoglou, S. (2010). Identifying a High Fraction of the Human Genome to be under Selective Constraint Using GERP++. *PLOS Computational Biology*, *6*(12), e1001025. https://doi.org/10.1371/JOURNAL.PCBI.1001025

Dedato, M., Robert, C., Taillon, J., Shafer, A., & Cote, S. (2022). *Demographic history and conservation genomics of caribou (Rangifer tarandus) in Québec* [Preprint]. Preprints. https://doi.org/10.22541/au.164165840.02956352/v1

DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, *30*(17), 4147–4154. https://doi.org/10.1111/mec.16051

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422–433. https://doi.org/10.1038/nrg.2016.58

Festa-Bianchet, M., & Côté, S. D. (2007). *Mountain Goats: Ecology, Behavior, and Conservation of an Alpine Ungulate*. Island Press.

Foote, A. D., Hooper, R., Alexander, A., Baird, R. W., Baker, C. S., Ballance, L., Barlow, J., Brownlow, A., Collins, T., Constantine, R., Dalla Rosa, L., Davison, N. J., Durban, J. W., Esteban, R., Excoffier, L., Martin, S. L. F., Forney, K. A., Gerrodette, T., Gilbert, M. T. P., … Morin, P. A. (2021). Runs of homozygosity in killer whale genomes provide a global record of demographic histories. *Molecular Ecology*, *30*(23), 6162–6177. https://doi.org/10.1111/mec.16137

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

Garcia-Dorado, A. (2015). On the consequences of ignoring purging on genetic recommendations for minimum viable population rules. *Heredity*, *115*(3), 185–187. https://doi.org/10.1038/hdy.2015.28

Garcia-Dorado, A., & Caballero, A. (2021). Neutral genetic diversity as a useful tool for conservation biology. *Conservation Genetics*, *22*(4), 541–545. https://doi.org/10.1007/s10592-021-01384-9

Gillespie, J. H. (2004). *Population Genetics: A Concise Guide*. JHU Press.

Glémin, S. (2003). How Are Deleterious Mutations Purged? Drift Versus Nonrandom Mating. *Evolution*, *57*(12), 2678–2687. https://doi.org/10.1111/j.0014-3820.2003.tb01512.x

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), 1001. https://doi.org/10.1038/s41467-020-14803-1

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., & Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, *9*(1), R7. https://doi.org/10.1186/gb-2008-9-1-r7

Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding Inbreeding Depression, Purging, and Genetic Rescue. *Trends in Ecology & Evolution*, *31*(12), 940–952. https://doi.org/10.1016/j.tree.2016.09.005

Huber, C. D., Kim, B. Y., & Lohmueller, K. E. (2020). Population genetic models of GERP scores suggest pervasive turnover of constrained sites across mammalian evolution. *PLoS Genetics*, *16*(5), e1008827. https://doi.org/10.1371/journal.pgen.1008827

Hunt, S. E., McLaren, W., Gil, L., Thormann, A., Schuilenburg, H., Sheppard, D., Parton, A., Armean, I. M., Trevanion, S. J., Flicek, P., & Cunningham, F. (2018). Ensembl variation resources. *Database: The Journal of Biological Databases and Curation*, *2018*. https://doi.org/10.1093/database/bay119

IUCN Standards and Petitions Committee. (2019). *Guidelines for Using the IUCN Red List Categories and Criteria* (No. 14; pp. 1–60). IUCN Red List.

Kardos, M., Armstrong, E. E., Fitzpatrick, S. W., Hauser, S., Hedrick, P. W., Miller, J. M., Tallmon, D. A., & Funk, W. C. (2021). The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences*, *118*(48), e2104642118. https://doi.org/10.1073/pnas.2104642118

Kardos, M., & Shafer, A. B. A. (2018). The Peril of Gene-Targeted Conservation. *Trends in Ecology & Evolution*, *33*(11), 827–839. https://doi.org/10.1016/j.tree.2018.08.011

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

Kauffman, M. J., Meacham, J. E., Hall, S., Steingisser, A. Y., Rudd, W. J., & Ostlind, E. (2018). *Wild Migrations: Atlas of Wyoming's Ungulates* (Vol. 31). University of Nebraska Press. https://muse.jhu.edu/article/789690

Kessler, C., Wootton, E., & Shafer, A. B. A. (2022). *Speciation without gene-flow in hybridizing deer*. https://doi.org/10.1101/2022.04.20.488928

Khan, A., Patel, K., Shukla, H., Viswanathan, A., van der Valk, T., Borthakur, U., Nigam, P., Zachariah, A., Jhala, Y. V., Kardos, M., & Ramakrishnan, U. (2021). Genomic evidence for inbreeding

depression and purging of deleterious genetic variation in Indian tigers. *Proceedings of the National Academy of Sciences*, *118*(49), e2023018118. https://doi.org/10.1073/pnas.2023018118

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

Kohn, M. H., Murphy, W. J., Ostrander, E. A., & Wayne, R. K. (2006). Genomics and conservation genetics. *Trends in Ecology & Evolution*, *21*(11), 629–637. https://doi.org/10.1016/J.TREE.2006.08.001

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics 2014 15:1*, *15*(1), 1–13. https://doi.org/10.1186/S12859-014-0356-4

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, *4*(7), 1073–1081. https://doi.org/10.1038/nprot.2009.86

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, *34*(7), 1812–1819. https://doi.org/10.1093/MOLBEV/MSX116

Kyriazis, C. C., Wayne, R. K., & Lohmueller, K. E. (2021). Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. *Evolution Letters*, *5*(1), 33–47. https://doi.org/10.1002/evl3.209

Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. Columbia University Press.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/BIOINFORMATICS/BTP324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/BIOINFORMATICS/BTP352

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. O., & Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*, *33*(20), 6494–6506. https://doi.org/10.1093/nar/gki937

Martchenko, D., Chikhi, R., & Shafer, A. B. A. (2020). Genome Assembly and Analysis of the North American Mountain Goat (Oreamnos americanus) Reveals Species-Level Responses to Extreme Environments. *G3: Genes, Genomes, Genetics*, *10*(2), 437–442. https://doi.org/10.1534/G3.119.400747

Martchenko, D., & Shafer, A. (2022). *Contrasting whole-genome and reduced representation sequencing for population demographic inference: An alpine mammal exampl*. https://doi.org/10.22541/au.164190604.45542461/v1

McDonald, J. S., & Miller, K. V. (2004). *A History of White-tailed Deer Restocking in the United States, 1878 to 2004*. Quality Deer Management Assoc.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology 2016 17:1*, *17*(1), 1–14. https://doi.org/10.1186/S13059-016-0974-4

Pascual, U., Adams, W. M., Díaz, S., Lele, S., Mace, G. M., & Turnhout, E. (2021). Biodiversity and the challenge of pluralism. *Nature Sustainability*, *4*(7), 567–572. https://doi.org/10.1038/s41893-021-00694-7

Peart, C. R., Tusso, S., Pophaly, S. D., Botero-Castro, F., Wu, C.-C., Aurioles-Gamboa, D., Baird, A. B., Bickham, J. W., Forcada, J., Galimberti, F., Gemmell, N. J., Hoffman, J. I., Kovacs, K. M., Kunnasranta, M., Lydersen, C., Nyman, T., de Oliveira, L. R., Orr, A. J., Sanvito, S., … Wolf, J. B. W. (2020). Determinants of genetic variation across eco-evolutionary scales in pinnipeds. *Nature Ecology & Evolution*, *4*(8), 1095–1104. https://doi.org/10.1038/s41559-020-1215-5

Petit-Marty, N., Vázquez-Luis, M., & Hendriks, I. E. (2021). Use of the nucleotide diversity in COI mitochondrial gene as an early diagnostic of conservation status of animal species. *Conservation Letters*, *14*(1), e12756. https://doi.org/10.1111/conl.12756

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/BIOINFORMATICS/BTQ033

Renaud, G., Hanghøj, K., Korneliussen, T. S., Willerslev, E., & Orlando, L. (2019). Joint Estimates of Heterozygosity and Runs of Homozygosity for Modern and Ancient Samples. *Genetics*, *212*(3), 587–614. https://doi.org/10.1534/GENETICS.119.302057

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.e4. https://doi.org/10.1016/j.cub.2018.08.066

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

Rogers, R. L., & Slatkin, M. (2017). Excess of genomic defects in a woolly mammoth on Wrangel island. *PLOS Genetics*, *13*(3), e1006601. https://doi.org/10.1371/journal.pgen.1006601

Schmidt, C., Hoban, S., Hunter, M., Paz-Vinas, I., & Garroway, C. J. (2022). *The IUCN Red List is not sufficient to protect genetic diversity*. EcoEvoRxiv. https://doi.org/10.32942/osf.io/hn4by

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., Colling, G., Dalén, L., Meester, L. D., Ekblom, R., Fawcett, K. D., Fior, S., Hajibabaei, M., Hill, J. A., Hoezel, A. R., Höglund, J., Jensen, E. L., Krause, J., Kristensen, T. N., … Zieliński, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*(2), 78–87. https://doi.org/10.1016/j.tree.2014.11.009

Stanke, M., Tzvetkova, A., & Morgenstern, B. (2006). AUGUSTUS at EGASP: Using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biology*, *7 Suppl 1*, S11.1-8. https://doi.org/10.1186/gb-2006-7-s1-s11

Streit, K., Hammacher, C., Zeller, A., & Hack, S. (2012). Sambamba: A Runtime System for Online Adaptive Parallelization. In M. O'Boyle (Ed.), *Compiler Construction* (pp. 240–243). Springer. https://doi.org/10.1007/978-3-642-28652-0_13

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, *123*(3), 585–595.

Teixeira, J. C., & Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. *Proceedings of the National Academy of Sciences*, *118*(10), e2015096118. https://doi.org/10.1073/pnas.2015096118

Väli, Ü., Einarsson, A., Waits, L., & Ellegren, H. (2008). To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology*, *17*(17), 3808–3817. https://doi.org/10.1111/j.1365-294X.2008.03876.x

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*, *43*(1), 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

van der Valk, T., Díez-del-Molino, D., Marques-Bonet, T., Guschanski, K., & Dalén, L. (2019). Historical Genomes Reveal the Genomic Consequences of Recent Population Decline in Eastern Gorillas. *Current Biology*□*: CB*, *29*(1), 165-170.e6. https://doi.org/10.1016/J.CUB.2018.11.055

Vaser, R., Adusumalli, S., Ngak Leng, S., Sikic, M., & Ng, P. C. (2015). SIFT missense predictions for genomes. *Nature Protocols*, *11*. https://doi.org/10.1038/nprot.2015.123

von Seth, J., Dussex, N., Díez-del-Molino, D., van der Valk, T., Kutschera, V. E., Kierczak, M., Steiner, C. C., Liu, S., Gilbert, M. T. P., Sinding, M. H. S., Prost, S., Guschanski, K., Nathan, S. K. S. S.,

Brace, S., Chan, Y. L., Wheat, C. W., Skoglund, P., Ryder, O. A., Goossens, B., … Dalén, L. (2021). Genomic insights into the conservation status of the world's last remaining Sumatran rhinoceros populations. *Nature Communications 2021 12:1*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-22386-8

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

Tables and Figures

Table 1: Table showing the number of genes and average number of deleterious mutations per species identified by GERP and SIFT. The number of genes was taken from the annotation of each species.

Species	Genes	SNPs in CEs (GERP)	SNPs in CEs with RS > 2 (GERP)	Deleterious mutations (SIFT)
Caribou	33,242	74	32	4,944
Mountain Goat	22,292	212	74	3,190
White-Tailed Deer	20,750	411	187	6,833

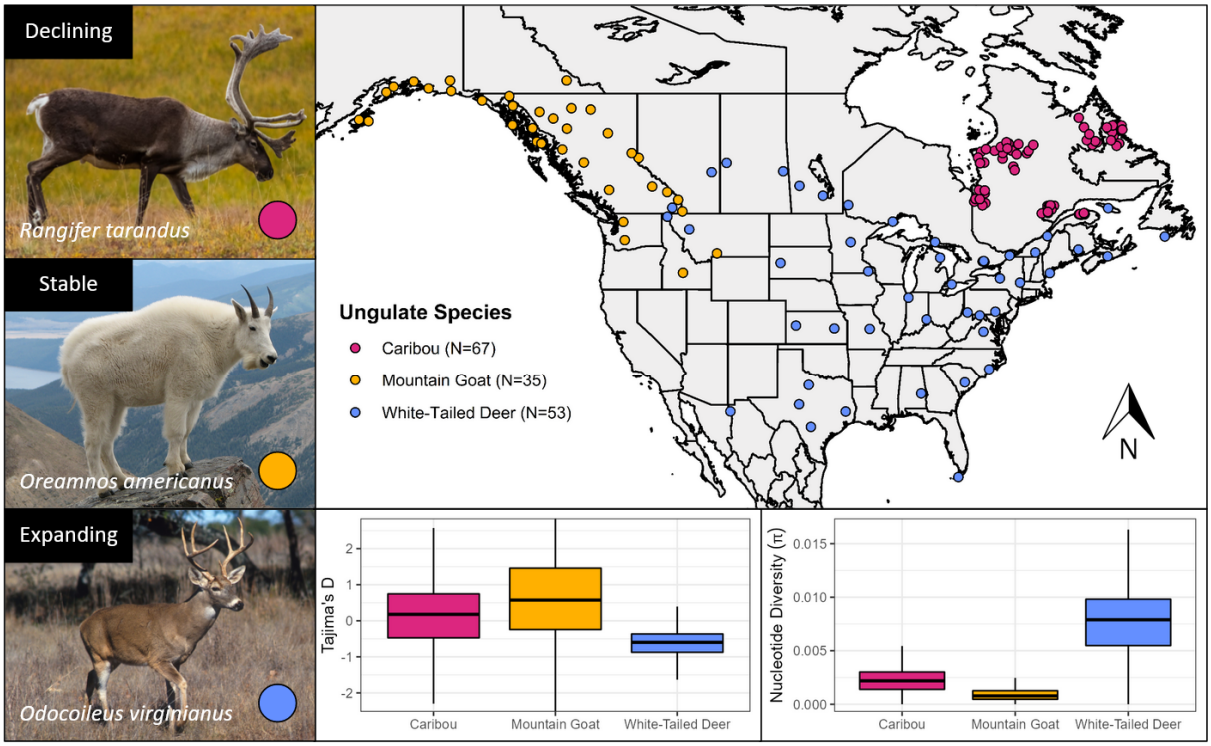

Figure 1: Sample locations and images of each species. (Red) *Rangifer tarandus* (N=67) (Yellow) *Oreamnos americanus* (N=35). (Blue) *Odocoileus virginianus* (N=53). Current demographic trends are outlined in the top left corner of each image. Boxplots showing Tajima's D and nucleotide diversity (π) per species are also pictured.

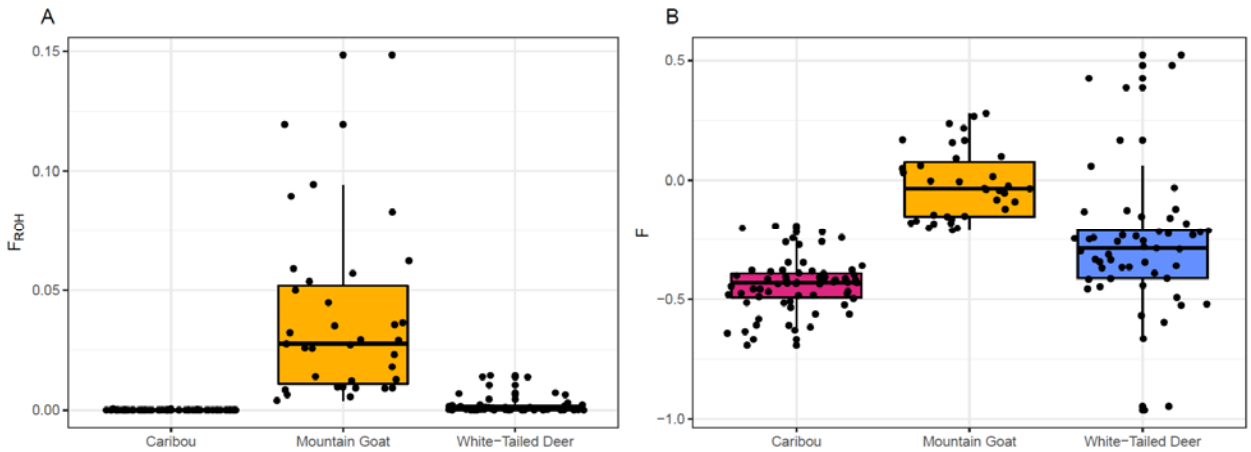

Figure 2: (A) Boxplot showing F_{ROH} as calculated by PLINK for all three species. Only ROHs greater than 100 Kb are included in each estimate. No ROHs were detected in caribou. (B) Boxplot showing the method-of-moments inbreeding coefficient (*F*) for each species as calculated by PLINK.

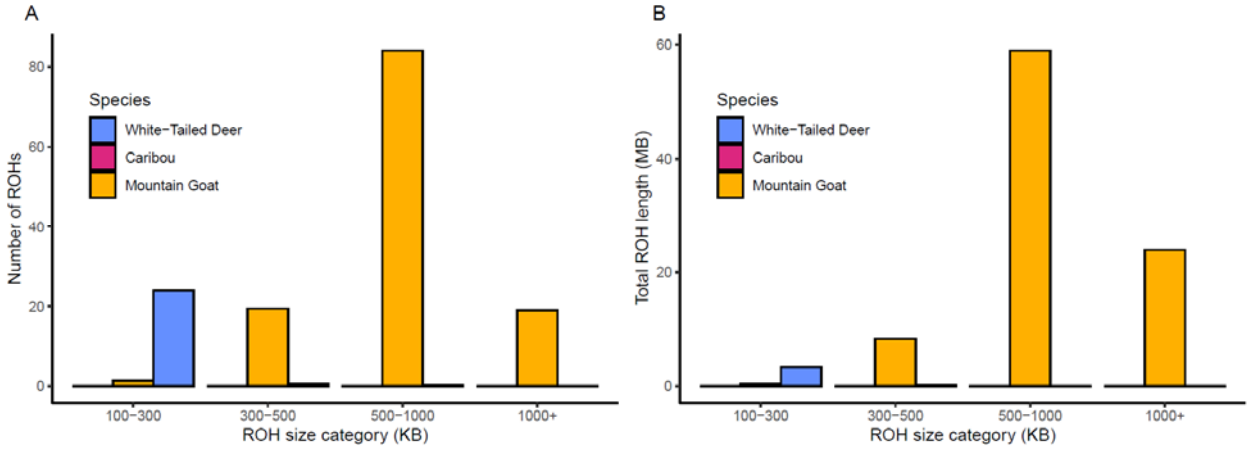

Figure 3: (A) Bar plot showing the total number of ROHs in each species stratified by ROH size category in Kb. No ROHs were detected in caribou. (B) Bar plot showing the total length (Mb) of all ROHs in each size category (Kb). No ROHs were detected in caribou.

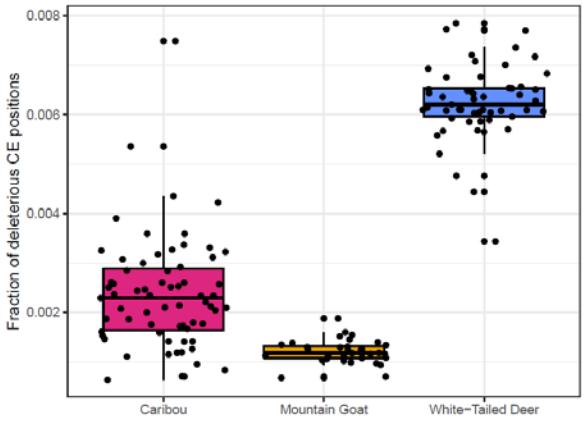

Figure 4: Boxplot showing the fraction of constrained positions with SNPs by species. Such SNPs are putatively deleterious, making this the fraction of positions that are deleterious in constrained elements (CEs).

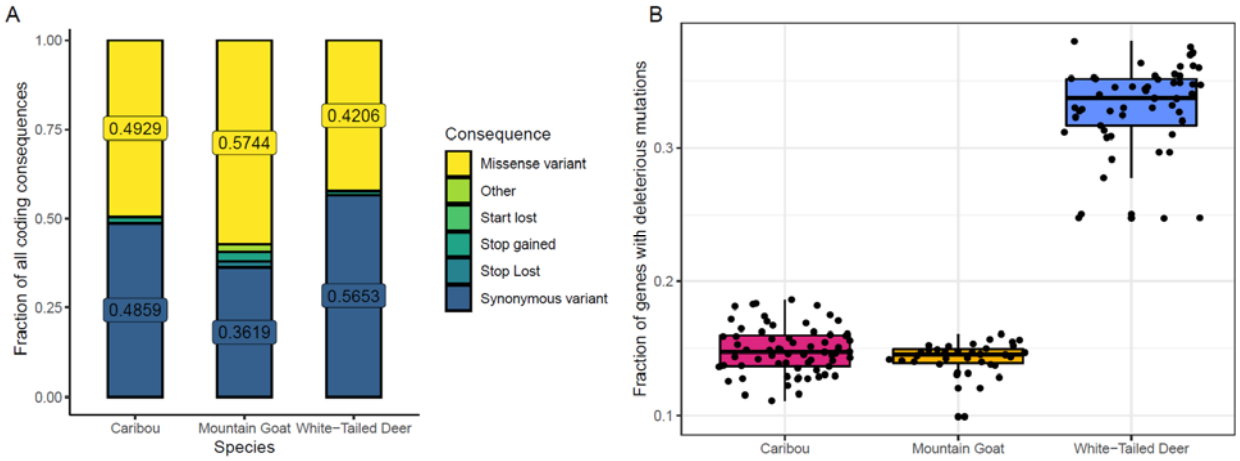

Figure 5: (A) Bar plot showing the fractions of each coding consequence identified by VEP in each species. (B) Boxplot showing the fraction of genes that possess deleterious mutations identified by SIFT.