CellPress
REVIEWS

Opinion

Quantifying Temporal Genomic Erosion in Endangered Species

David Díez-del-Molino,[1] Fatima Sánchez-Barreiro,[2] Ian Barnes,[3] M. Thomas P. Gilbert,[2,4] and Love Dalén[1,*]

Many species have undergone dramatic population size declines over the past centuries. Although stochastic genetic processes during and after such declines are thought to elevate the risk of extinction, comparative analyses of genomic data from several endangered species suggest little concordance between genome-wide diversity and current population sizes. This is likely because species-specific life-history traits and ancient bottlenecks overshadow the genetic effect of recent demographic declines. Therefore, we advocate that temporal sampling of genomic data provides a more accurate approach to quantify genetic threats in endangered species. Specifically, genomic data from predecline museum specimens will provide valuable baseline data that enable accurate estimation of recent decreases in genome-wide diversity, increases in inbreeding levels, and accumulation of deleterious genetic variation.

Highlights

The small population size of many endangered species makes them vulnerable to genetic threats, such as inbreeding depression, loss of genetic variation, and accumulation of deleterious mutations

Present-day genome-wide diversity is a poor predictor of population size and conservation status in endangered species, making interspecific comparisons of genomic data inadequate indicators of extinction risk.

Historical specimens in museum collections include samples that pre-date the demographic declines that have occurred in recent centuries and, thus, can provide baseline levels of diversity, inbreeding, and genetic load.

Temporal genomic indices based on comparisons of historical and present-day samples are powerful tools to estimate genomic erosion, and could be integrated with other IUCN Red List criteria to assess threat levels in endangered species.

Genetic Threats to Small Populations

The population sizes of many wild organisms have gone through dramatic declines over the past 200 years as a consequence of human activities [1]. This has raised the concern that processes related to small population size might lead to an increased risk of extinction [2]. The underlying mechanisms include not only stochastic demographic and environmental events [3], but also genetic processes, such as increased **genetic drift** and **inbreeding** levels (see Glossary) [4]. For example, strong genetic drift can lead to a loss in standing genetic variation, reducing the **adaptive potential** of the population [5]. Moreover, increased genetic drift and inbreeding can lead to **inbreeding depression** through increased exposure of recessive deleterious alleles in homozygotes, as well as an increase in homozygosity at loci in which heterozygous genotypes have a fitness advantage [6].

Additionally, theoretical work suggests that genetic drift in small populations can become so strong that the ability of **purifying selection** to remove detrimental alleles is reduced [7], causing accumulation and fixation of deleterious variants throughout genomes. This accumulation of deleterious alleles can result in negative population growth, which in turn would lead to even higher genetic drift and more deleterious alleles becoming fixed, a process that ultimately might result in extinction [8].

There is a growing body of empirical evidence from both laboratory and wild settings showing that high inbreeding levels and loss of genetic variation can impact population viability [9]. However, in populations where deleterious alleles have become fixed, or in cases where all individuals are equally inbred, detecting decreased fitness via inbreeding depression using pedigree data or heterozygosity-fitness correlations is not possible [10]. While computer simulations have shown that inbreeding depression can elevate extinction risk (e.g., [11]),

[1]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, 10405 Stockholm, Sweden
[2]Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark
[3]Natural History Museum, Cromwell Road, SW7 5BD London, UK
[4]Norwegian University of Science and Technology, University Museum, 7491 Trondheim, Norway

*Correspondence: love.dalen@nrm.se (L. Dalén).

 https://doi.org/10.1016/j.tree.2017.12.002

comprehensive evidence from threatened wild populations is lacking, as is evidence that processes, such as the accumulation of deleterious alleles, have a direct impact on endangered species. Therefore, given the potential genetic threats discussed above, there is a need to better understand to what extent recent demographic declines have led to losses in genetic variation, as well as increased inbreeding levels and **genetic load**, in endangered species [12].

An Unfulfilled Promise of Modern Genomics

Conservation geneticists have traditionally used genetic markers, such as mitochondrial DNA and autosomal microsatellite markers, to survey genetic variation (e.g., [13–15]). Moreover, unless pedigree data have been available, microsatellites have been used to estimate inbreeding levels (e.g., [16]). However, the suitability of such markers for population genetics purposes has been repeatedly questioned, because they have been found to be underpowered and to render contradictory results [17–20].

With the advent of high-throughput sequencing techniques, it is now feasible to sample genome-wide diversity at the population level [21]. As sequencing costs decrease, *de novo* assembled and annotated **reference genomes** are being generated for a rapidly growing number of non-model organisms and these, in turn, provide the foundation for population-scale resequencing projects [22–26]. Such **whole-genome sequencing** data now enable conservation geneticists to obtain accurate estimates of present-day levels of genome-wide diversity and inbreeding levels, as well as detecting the presence of deleterious alleles [5,27–30]. Furthermore, genome sequencing also makes it possible to quantify copy number variants (CNVs), deletions, and even larger scale rearrangements. Such variants have long been known to hold significant relevance for genomic and phenotypic functions, although, until recently, most of our understanding of the extent of their roles was limited to the context of the evolution, health, and disease of humans and other model organisms [31,32].

Nonetheless, despite the apparent wealth of possibilities offered by the application of genomics to conservation, and the initial enthusiasm for sequencing large numbers of individuals from present-day populations, there is a risk that the lack of an appropriate strategy might render genomics uninformative for conservation and management [27]. Indeed, several recent studies that have estimated genome-wide diversity in endangered species [21,25,33] show that, even though genome-wide diversity estimates are correlated with current population sizes (Figure 1A), this correlation is rather weak ($r^2 = 0.29$) and there is no relationship between genome-wide diversity estimates and the IUCN Red List status of a species (Figure 1B). Similarly, there is no apparent correlation between genome-based estimates of inbreeding or genetic load and current population sizes in endangered taxa [25]. Instead, it appears that factors other than recent demographic changes, such as ancient **population bottlenecks** [34] and species-specific life-history traits [35], are more important in influencing genome-wide diversity in wild organisms, and that these, consequently, overshadow the effects of recent declines in population size. Taken together, these observations suggest that interspecific comparisons of contemporary genome-wide diversity are poor proxies for population size and conservation status in wild organisms and provide little information on the extent of genome erosion that endangered species have been subjected to as they approach the brink of extinction.

Conservation Paleogenomics: Establishing Predecline Baselines

We advocate that analyses of temporally sampled genomic data can provide a more accurate approach to quantify the genetic threats that endangered species face, as opposed to relying exclusively on present-day genomic data. Specifically, comparisons of present-day genomic

Glossary

Adaptive genetic diversity: fraction of the genetic variation in a population that, given a particular set of environmental conditions, has a positive effect on the fitness of the individuals carrying it and, thus, is susceptible to being driven to high frequencies in the population by positive selection.
Adaptive potential: also termed evolvability or evolutionary potential; capability of a population to respond to selective forces as environmental conditions change.
Deleterious or detrimental genetic variation: fraction of the genetic variation in a population that has, or could have, a negative impact on the fitness of the individuals that carry it.
Genetic drift: stochastic changes in allele frequencies across generations of a population of finite size due to the random sampling of parental alleles.
Genetic load: proportion of mildly deleterious genetic variants accumulated across the genome. When random genetic drift is a strong force (i.e., in small populations), mildly deleterious genetic variants are not as likely to be purged, but rather fixed by drift, thus increasing the genetic load.
Genome-wide heterozygosity: proportion of sites across the genome that are polymorphic between the genome copies of a diploid or polyploid organism.
Genomic erosion: collective term that includes the entire array of negative genetic changes caused by random processes in small populations, such as loss of genome-wide diversity, alterations in genomic structural variants (e.g., deletions), as well as the accumulation of runs of homozygosity and deleterious mutations.
Identical by descent: status of two identical alleles in a diploid genome that are so because they derive from a common ancestral allele.
Inbreeding: the mating of relatives, where the inbreeding coefficient (F) is defined as the probability that two alleles in an individual are identical by descent.
Inbreeding depression: reduction in the viability of a population due to

data with that from museum specimens pre-dating demographic declines enable the direct quantification of recent changes in genetic parameters on a genome-wide scale [36–38]. Such changes are particularly relevant in a conservation context, because populations that have gone through rapid declines in size are thought to be more prone to inbreeding depression than are populations that have been small for a long time [5]. The parameters that can be estimated using a temporal genomic approach include not only changes in **neutral genetic diversity** and **adaptive genetic diversity**, but also changes in inbreeding levels (see below and Box 1). Moreover, this approach can also be used to obtain accurate estimates of whether, and to what extent, **deleterious genetic variation** has accumulated through time.

The analysis of serially sampled population genomic data is routinely being used in ancient DNA studies, on samples that are several thousands of years old (e.g., [37,39,40]). Moreover, serial sampling of museum specimens using genetic markers, such as mitochondrial DNA and microsatellite loci, was introduced during the early 1990s, and has since been applied to a range of endangered taxa [41,42]. However, we propose that a temporal approach using genomic data is already possible, can provide valuable information that cannot be obtained using traditional marker types, and, thus, needs to be exploited in conservation biology.

The extensive demographic declines that most currently endangered species have gone through occurred during the past few centuries [1]. Consequently, many museum collections contain specimens that pre-date the onset of these demographic declines. This means that natural-history collections constitute archives of the past that are valuable for conservation genomics, since these collections can be used to establish baseline levels of genome-wide diversity. Indeed, several recent studies have shown that museum specimens stored at ambient temperatures, even for several hundreds of years, can contain relatively high proportions of endogenous DNA [38,43–47], particularly if they are from tissues that were preserved in a manner that prevented microbe-driven putrefaction (e.g., drying or tanning, pickling in alcohol, or stripping off the soft tissues shortly after time of death, in the case of bones). The recovered DNA is often highly fragmented, with the retrieved sequences that can be unambiguously aligned to the reference genomes normally ranging from 25 to a few hundred base pairs (bp; typically peaking at 40–80 bp [48]; Box 2). The high-throughput sequencing platforms currently favored by paleogenomicists, such as the Illumina, IonTorrent, and BGISeq series, are ideally placed to recover the full sequence information from these types of sample.

Applying a shotgun sequencing approach to historical museum samples could also be used to retrieve an array of additional information about evolutionary and ecological changes through time in endangered species, such as changes in the microbiome, associated pathogens, as well as methylation patterns of a species (see Outstanding Questions). Moreover, recovering DNA from dental calculus [49] or paleofeces [50] could provide a powerful way to examine changes in diet that occurred in conjunction with demographic declines in endangered species.

Quantifying Temporal Genomic Erosion

There are two important challenges to accurately and effectively infer temporal genomic changes in endangered species that are relevant in a conservation perspective. First, to reconstruct multiple genomes from both modern and historical samples, a reference genome is needed. To maximize quality, the reference should ideally be a *de novo* assembly from the same species. Alternatively, a reference from a closely related species could be used, although, in that case, the divergence time also needs to be considered. This is because fewer sequence reads will be retrieved the more distant the reference is, and this can lead to biases in the resulting genome [51,52]. Second, there is a lack of established methods to accurately infer

either inbreeding or genetic drift. Both processes lead to an increase in homozygosity at the population level, which will in turn expose detrimental alleles, and augment the proportion of homozygosity at loci where the heterozygote has a fitness advantage.

Loss-of-function (LoF) variants: alterations of the DNA sequence with pervasive effects on the expression and regulatory mechanisms of the genic regions of a genome.

Neutral genetic diversity: fraction of the genetic variation in a population that does not have any effect on the fitness of the individuals that bear it, and whose frequency is, therefore, not determined directly by selective forces.

Paleogenomics: scientific field devoted to the genomic analysis of historical samples using ancient DNA technology.

Population bottleneck: severe reduction in population size due to extrinsic, nongenetic factors.

Purging: enhanced form of purifying selection that acts on small populations in which increased levels of homozygosity expose recessive deleterious alleles.

Purifying selection: selective force that eliminates detrimental alleles from a population.

Reference genome: consensus DNA sequence stored as a digital dataset that represents, with a relatively high level of completeness, the actual genome of a given species.

Runs of homozygosity: stretches of genomic sequence that are identical at every position on the two genomic copies of a diploid organism. Inbreeding and, at a more localized scale, selection can generate runs of homozygosity, which are in turn disrupted over time through mutation and recombination.

Whole-genome sequencing: acquisition of genomic data that span as much of the entire genome sequence of an organism as possible by means of sequencing techniques.

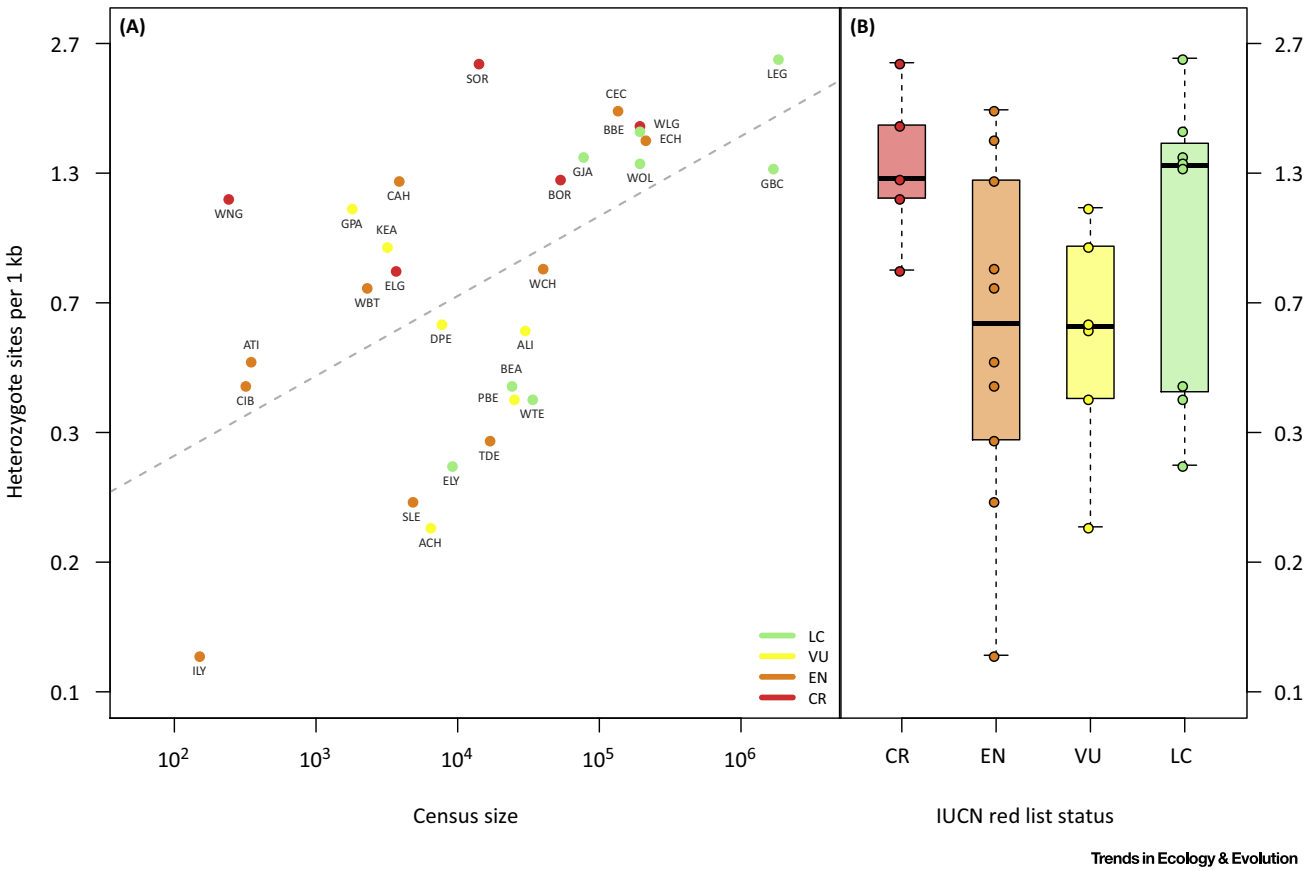

Figure 1. Relationship between Genome-Wide Heterozygosity, Census Population Size, and Conservation Status in Endangered Species. Approximate census population sizes and threat categories (in different colors), as extracted from the IUCN Red List (www.iucnredlist.org/). All genome-wide heterozygosity values were obtained from whole-genome sequencing studies [22,25,26,33,73–75]. (A) Relationship between census population size and present-day genome-wide heterozygosity (r^2 = 0.29, F = 12.51, P = 0.0015). (B) Genome-wide heterozygosity levels for species with different IUCN Red List status. Heterozygosity levels for different Red List categories were not significantly different (ANOVA, F = 2.06, P = 0.14). The taxa used in the comparisons are: WLG, Western lowland gorilla, *Gorilla gorilla gorilla*; WNG, Western Nigerian gorilla, *Gorilla gorilla diehli*; ELG, Eastern lowland (Grauer's) gorilla, *Gorilla beringei graueri*; SOR, Sumatran orangutan, *Pongo abelii*; CAH, Cameroon chimpanzee, *Pan troglodytes ellioti*; WCH, Western chimpanzee, *Pan troglodytes verus*; ECH, Eastern chimpanzee, *Pan troglodytes schweinfurthii*; CEC, Central chimpanzee, *Pan troglodytes troglodytes*; BOR, Bornean orangutan, *Pongo pygmaeus*; ALI, African lion, *Panthera leo*; ATI, Amur tiger, *Panthera tigris altaica*; TDE, Tasmanian devil, *Sarcophilus harrisii*; SLE, snow leopard, *Panthera uncia*; GPA, giant panda, *Ailuropoda melanoleuca*; WBT, white Bengal tiger, *Panthera tigris tigris*; PBE, polar bear, *Ursus maritimus*; BBE, ABC brown bear, *Ursus arctos*; GJA, golden jackal, *Canis aureus*; WOL, wolf, *Canis lupus lupus*; CIB, crested ibis, *Nipponia nippon*; LEG, little egret, *Egretta garzetta*; DPE, Dalmatian pelican, *Pelecanus crispus*; GBC, great black cormorant, *Phalacrocorax carbo*; KEA, kea, *Nestor notabilis*; BEA, bald eagle, *Haliaeetus leucocephalus*; WTE, white-tailed eagle, *Haliaeetus albicilla*; ILY, Iberian lynx, *Lynx pardinus*; ELY, Eurasian lynx, *Lynx lynx*; ACH, African cheetah, *Acinonyx jubatus*.

temporal genomic parameters of conservation interest, likely because it has only recently become possible to generate such data. For example, current methods to detect microevolutionary processes, such as natural selection and genetic drift, are primarily based on inference using modern samples [53,54] and, thus, are unable to leverage the power of serial sampling to capture the specific dynamics of changes in both allele frequencies and genome architecture.

To tackle this issue, there are several genomic parameters that we believe are relevant in the context of endangered species and that can be easily incorporated into a temporal framework provided that genomes from at least two different time points are available (Box 1). This would enable the calculation of delta estimators that represent temporal changes in, for example, heterozygosity (Δh), inbreeding (ΔF), genetic load (ΔL), and genomic deletions (Δd).

Box 1. Genomic Erosion Methods and Delta Estimators

Individual **genome-wide heterozygosity** (h) is routinely estimated in genome sequencing studies using genotype hard calls. After variant calling (e.g., GATK [76]), h can be expressed as the number of heterozygous sites divided by the total number of genotyped sites. More elaborate methods make use of likelihood estimates of the population mutation parameter θ (mlRho [37,77]) or of genotype probabilities to estimate Waterson's theta θ_W (ANGSD [78]). Emerging approaches to estimate h can deal with low-quality data, such as low coverage and small fragment size [79], and DNA damage [80]. Genomic **runs of homozygosity** (ROHs) can be detected in sliding windows using PLINK [81], Watterson's theta (θ_W) and change-point analyses [82], with a LOD estimator [59], or using the TMRCAs estimated by PSMC (e.g., [37]). New methods, such as BCFTOOLS/RoH [83] and H3M2 [57], use a more-refined Hidden Markov Model approach to identify ROHs. To detect deleterious variants, genomes can be annotated with tools such as ANNOVAR [84], or lifted over from a closely related species. Tools such as VEP [85] can be used to predict deleterious variant consequences (i.e., [24]). Using a copy number variation strategy, genomic deletions can be identified as stretches of the genome in which the per site coverage is consistently lower than the genome average using a change-point analysis algorithm [82].

Regardless of what method is used, each index can easily be transformed into a conservation-relevant delta estimator that incorporates the temporal dimension of the change, by subtracting the baseline value of the predecline samples from the present-day ones (see Figure 2 in the main text). Thus, following theoretical predictions [8], endangered species that have suffered population declines would typically show negative values of Δh, implying net loss of genome-wide diversity through time; positive ΔF, indicating an increase on the inbreeding levels in recent generations; positive ΔL, suggesting increased genetic load; and positive Δd, because the frequency of deletions is also expected to accumulate in small populations [63].

Genome-Wide Heterozygosity

The change in genome-wide diversity caused by demographic declines provides a good proxy for the loss of adaptive potential that an endangered species has been subjected to. Accurate information about the genome-wide diversity of a population can be estimated by, for example, computing heterozygosity (h) obtained from whole-genome analyses of single diploid individuals [55]. However, a relevant drawback is that uneven coverage, presence of errors characteristic of shotgun sequencing experiments, and postmortem damage (PMD) patterns characteristic of historical and ancient DNA-sequencing experiments [56] can hinder the ability to retrieve reliable heterozygosity estimations. Therefore, generating less biased estimates of heterozygosity from historical samples might require relatively high genome coverage (typically $>10\times$).

Runs of Homozygosity

The level of inbreeding in a population can be assessed by estimating the amount of within-individual runs of homozygosity (ROHs). At the most fundamental level, ROHs can be defined as chromosomal stretches of a diploid genome that are in a homozygous state [57]. The total number and cumulative length of ROHs can be used to estimate individual inbreeding coefficients (F [58]). Moreover, the abundance and size distribution of ROHs provide additional information about the evolutionary history of a population [59]. While small ROHs (<200 kb) probably constitute remnants of ancient inbreeding events where the resulting ROHs have been fragmented due to recombination, longer ROHs represent stretches of DNA that are directly **identical by descent** and can be used to assess the levels of recent inbreeding in the population that the individual was sampled from [28]. In analyses comprising only modern data, small to intermediate-sized ROHs are typically disregarded since it is difficult to resolve whether they are a result of background relatedness or demographic bottlenecks from a distant past [59]. However, since ROHs caused by ancient bottlenecks would also be present in the baseline historical samples, accumulation of short ROHs resulting from recent (postdecline) inbreeding between distant relatives can be reliably identified [37] and, therefore, should also be incorporated in ΔF calculations (Box 1).

Box 2. The Trouble with Ancient DNA

Although the 'ancient' part of ancient DNA has been poorly defined since inception of the field during the early 1980s, our insights into how DNA degrades post mortem have been considerably assisted thanks to the development of high-throughput sequencing platforms that quantify the characteristics of ancient DNA [86–88]. Data from ancient DNA studies suggest that genetic material from specimens that were not explicitly stored with DNA extraction in mind will tend to be degraded in similar ways [89]. Thus, work on DNA from both 'archival' collections assembled in museums and herbaria over the past three centuries, and 'excavated' collections from paleontological and archaeological sites, benefits from the same specialized laboratory methods.

PCR-Sanger sequencing protocols suggested that ancient DNA was rarely longer than 200-or-so base pairs, was generally co-extracted with compounds that would inhibit downstream enzymatic reactions, and was chemically modified in a way that could confuse sequence recovery via the hydrolytic decay of cytosine, leading to its replacement by thymine in the resulting sequence [90,91]. Analysis of ancient DNA using high-throughput protocols provides a more balanced picture, with mean sequence lengths typically well under 100 base pairs, an exponential distribution of fragmentation, and low proportions of endogenous DNA relative to bacteria (typically, although not exclusively, a problem for excavated specimens) [48]. In addition, the small quantities of starting DNA necessitate multiple cycles of PCR during the library build process, reducing library complexity and, therefore, genome coverage and the ability to detect heterozygous sites.

Despite these problems, it has been possible to recover genomes and genome-scale datasets from multiple species using specimens of tens to tens of thousands of years in age [37,92,93]. Chief among these are **paleogenomic** studies of humans, where researchers benefit from the availability of multiple, high-quality genomes to map to, and huge datasets of genomic polymorphisms to compare their ancient data with. However, methodological developments, frequently driven by work on the Neanderthal genome, have facilitated the analysis of other animal, plant, and even pathogen species. These ancient DNA-specific methods include the use of skeletal elements with higher endogenous content, the use of enzymatic pretreatments to reduce the impact of DNA damage, and the development of highly efficient library-build processes [56,94,95].

Genetic Load

Loss-of-function (LoF) variants, such as stop-gain, stop-loss, splice site, and frame-shift mutations, are predicted to disrupt gene function and, hence, are expected to substantially reduce individual fitness and population viability [8]. Thus, even though unambiguously detecting and correctly classifying these variants can be challenging, because it requires functional tests that are not always possible in non-model species, quantifying their relative genetic burden in endangered species is of fundamental conservation interest [60].

The ratio between the number of LoF variants and the number of synonymous variants should be informative of the genetic load (L) per individual, relative to the amount of variation it has [60]. Thus, comparison of data from pre- and postdecline specimens would provide an estimate of how the genetic load of a population has changed through time. Note that this estimator would be positive if there has been a net accumulation of deleterious mutations in the contemporary population, and negative if purifying selection has reduced the overall genetic load (e.g., [25]). Therefore, estimating ΔL can help determine how common **purging** is in small endangered populations, which is an important question in conservation biology [61]. In addition, even if purging is found to have occurred in an endangered species (e.g., [24]), a subset of the deleterious mutations might nonetheless become fixed as a result of genetic drift. Identifying such fixed LoF mutations, which could be facilitated by temporal sampling, can be essential for conservation.

Genomic Deletions

Even though human populations are known to be polymorphic for tens of thousands of genomic deletions ranging in size from a few hundred base pairs to several megabases [62], these structural variants have generally been overlooked in conservation genetics. A recent comparison of two mammoth genomes [37] indicated that the specimen dated close to

the extinction of the species had an accumulation of homozygous deletions (d), including an elevated number of deleted exons [63]. This suggests that deletions of functionally important genetic variation might constitute a previously unknown genetic threat to endangered species. Thus, examining to what extent deletions, including those that might directly affect genes, have increased in frequency, or even become fixed as a consequence of recent population declines, should be of high conservation interest. If functional annotations for the studied species are available, then it would also be possible to estimate to what extent deletions that affect functionally important genes, or enriched functional categories [64], have increased in frequency as a consequence of population declines.

However, one important consideration when quantifying changes in the frequency of genomic deletions by comparing historical predecline and contemporary samples is that the reference genome to which sequencing reads from all the samples are mapped cannot belong to the present-day population. This is because such a reference genome would already have the deletions one is interested in identifying, hence biasing the comparisons. Consequently, unless a *de novo* genome assembly can be obtained from a predecline specimen, which, in most cases, is unlikely given that high-molecular-weight DNA is required, it will be necessary to instead map the resequencing data from both modern and historical samples to a reference genome from a closely related species.

Concluding Remarks and Practical Implications for Conservation

We have here discussed the value of incorporating temporal genomic data in conservation. Levels of genome-wide diversity and deleterious genetic variants identified in present-day endangered species provide only limited information about their conservation status. Instead, we argue that delta indices obtained by integrating estimates from both present-day samples and samples from museum collections that pre-date recent declines in population size provide a more accurate quantification of the temporal **genomic erosion** that species have been subjected to as they have approached the brink of extinction. This strategy is already possible, since the cost of generating genomic data has dropped significantly in recent years [27], to the extent that a high-coverage mammalian genome can now be sequenced for less than US$1000. We have proposed four indices that can be used to quantify genomic erosion using historical samples to estimate the baseline levels. As genomic technologies improve, for example enabling imputation from phased reference data in non-model organisms, additional indices, such as estimating haplotype diversity by treating large chromosomal stretches as alleles, could provide more powerful ways to estimate losses of genetic diversity.

A hard reality faced by conservation practitioners is that there are too many species under serious threat and, with too few funds available, the resources need to be used wisely [65]. In this regard, a key contribution is that, by identifying whether modern genomic profiles are significantly different to that of past populations, one can begin to prioritize species (or subpopulations) for which intervention might warrant alleviating genetic problems. For example, regardless of whether the genome of a particular species carries signals classically linked to poor population health, such as elevated levels of putatively deleterious mutations, if historic data indicate that this is not a recent phenomenon, then this might imply that these mutations have no significant conservation relevance. By contrast, should the historic data provide good evidence for major changes in genetic load over a short timespan, then we can not only use this to propose testable hypotheses to identify the function of putatively deleterious mutations (e.g., [66]), but also use those that likely are deleterious as measurable indicators of individual fitness as well as species viability. In addition, this information could be used as a monitoring tool with

Outstanding Questions

Which types of historical samples are best for genomic analysis? That is, how is DNA degradation in museum samples affected by different preservation conditions, such as temperature, humidity, and mode of storage (e.g., dry, ethanol, or formalin), and how do these parameters interact with the types of material typically held in museum collections (e.g., skin, hair, bone, feathers, or teeth)?

How should different levels of genome erosion be translated into specific IUCN Red List criteria? For example, at what magnitude of change in Δh or ΔL should the Red List status of a species be changed from Endangered to Critically Endangered?

To what extent have past extinctions been associated with losses of genetic variation, inbreeding, and accumulation of deleterious mutations before that species disappeared?

How have the microbiomes and epigenomes of endangered taxa been affected by the demographic declines that have occurred as these species have approached the brink of extinction?

How have recent changes in selection pressure, for example through increased levels of environmental toxins, overharvesting, climate change, or habitat modification, affected the genomic architecture in endangered species during the past 100–200 years?

which to measure ongoing changes in threatened species [67,68] and, as such, even be used to audit the success of ongoing interventions.

There is also a potential role for a more nuanced understanding of exactly which genomic indicators are relevant for particular species, with regards to guiding the actions of applied conservation measures (Figure 2). In particular, targeted translocation and breeding efforts can be approached in a manner that is more subtle than the current strategy of simply maximizing genetic diversity. For example, if historic evidence is used to identify putative LoF alleles that have increased in frequency in conjunction with a population collapse, this would provide a strong indication that these alleles have a negative impact on fitness. Conservation managers could use such information on specific deleterious alleles to select individuals suitable for translocation or *ex situ* management [69]. They could also take steps towards removing these deleterious alleles from the population (Figure 2), by means of, for example, selective breeding and translocations [70], or possibly through genome editing [71].

Intraspecific estimates of genomic erosion, such as the delta estimators discussed here, can provide quantifiable measures of extinction risk that are comparable across taxonomic groups. As such, these delta estimators would be suitable for integration with other criteria currently being used to assign threat categories in the IUCN Red List. For instance, genomic data showing that a species has lost a significant amount of genome-wide diversity and, thus, adaptive potential, and/or has gained a large number of LoF mutations over the past 100 years could provide justification for changing its conservation status to a higher threat category (Figure 2). As such, the ghost of genetic diversity past [72], which is now accessible through genomic analyses of historical museum specimens, can help prevent future extinctions through more accurate threat assessments, and could also serve as a benchmark for the restoration of biodiversity at its most fundamental level.

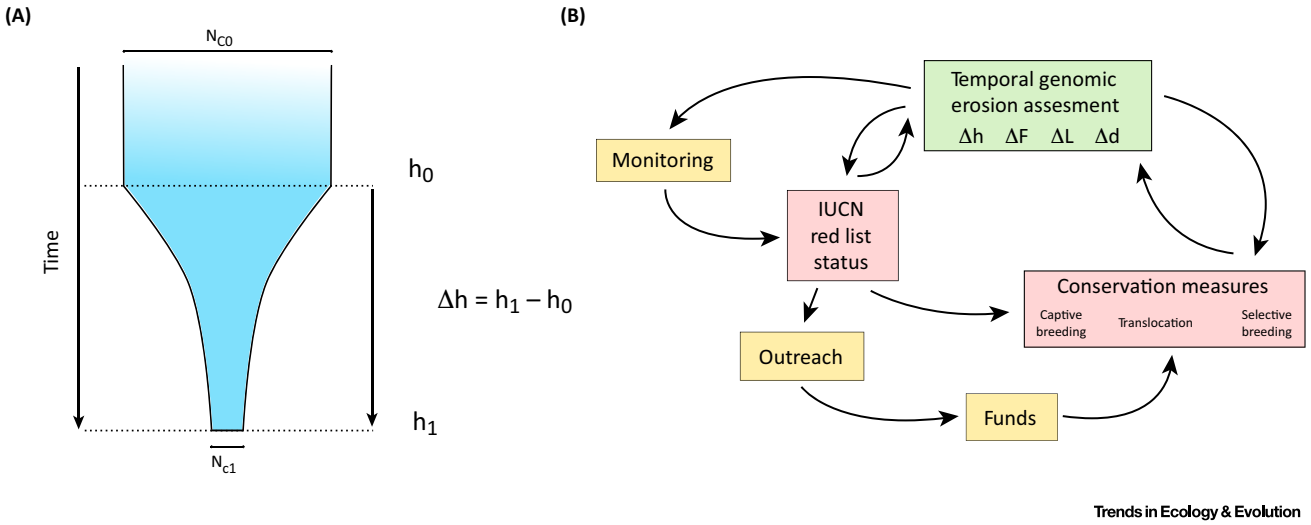

Figure 2. Quantifying Temporal Genomic Erosion and Possible Conservation Implications. (A) Schematic view of changes in the census population size (N_C) of an endangered species through time, showing a dramatic decline from N_{C0} to N_{C1}. In this example, h_0 represents the genome-wide heterozygosity of individuals from museum collections sampled before the population decline, whereas h_1 represents heterozygosity in present-day individuals. The ability of delta estimators to capture the temporal dimension of the change is indicated (see Box 1 in the main text). (B) Overview of possible implications of the proposed temporal approach. Delta estimators of genomic erosion might be used to assign IUCN Red List threat-level categories in endangered species, and can also be used for monitoring purposes, as well as to better inform tailored conservation and management measures for species at the brink of extinction.

Acknowledgments

The authors would like to thank four anonymous reviewers for constructive comments on an earlier version of the manuscript. The authors also acknowledge funding from FORMAS grant 2015-676 and VR grant 2012-3869 (to L.D.), NERC grant NE/J010480/1 (to I.B.), and the ERC-COG grant 'Extinction Genomics' (to M.T.P.G.).

References

1. Dirzo, R. *et al.* (2014) Defaunation in the Anthropocene. *Science* 345, 401–406
2. Caughley, G. (1994) Directions in conservation biology. *J. Anim. Ecol.* 63, 215–244
3. Melbourne, B.A. and Hastings, A. (2008) Extinction risk depends strongly on factors contributing to stochasticity. *Nature* 454, 100–103
4. Frankham, R. (2005) Resolving the genetic paradox in invasive species. *Heredity* 94, 385
5. Kohn, M.H. *et al.* (2006) Genomics and conservation genetics. *Trends Ecol. Evol.* 21, 629–637
6. Charlesworth, B. and Charlesworth, D. (1999) The genetic basis of inbreeding depression. *Genet. Res.* 74, 329–340
7. Kimura, M. (1957) Some problems of stochastic processes in genetics. *Ann. Math. Stat.* 28, 882–901
8. Lynch, M. *et al.* (1995) Mutation accumulation and the extinction of small populations. *Am. Nat.* 146, 489–518
9. Neaves, L.E. *et al.* (2015) The fitness consequences of inbreeding in natural populations and their implications for species conservation – a systematic map. *Environ. Evid.* 4, 5
10. Keller, L.F. and Waller, D.M. (2002) Inbreeding effects in wild populations. *Trends Ecol. Evol.* 17, 230–241
11. O'Grady, J.J. *et al.* (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biol. Conserv.* 133, 42–51
12. Frankham, R. (2005) Genetics and extinction. *Biol. Conserv.* 126, 131–140
13. Harley, E.H. *et al.* (2005) Genetic variation and population structure in remnant populations of black rhinoceros, *Diceros bicornis*, in Africa. *Mol . Ecol.* 14, 2981–2990
14. Gottelli, D. *et al.* (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Mol. Ecol.* 3, 301–312
15. Chang, Y.L. *et al.* (2006) Bayesian estimation of the timing and severity of a population bottleneck from ancient DNA. *PLoS Genet.* 2, e59
16. White, K.L. *et al.* (2015) Evidence of inbreeding depression in the critically endangered parrot, the kakapo. *Anim. Conserv.* 18, 341–347
17. Galtier, N. *et al.* (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol. Ecol.* 18, 4541–4550
18. Moritz, C. (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Mol. Ecol.* 3, 401–411
19. Väli, U. *et al.* (2008) To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Mol. Ecol.* 17, 3808–3817
20. Balloux, F. *et al.* (2004) Does heterozygosity estimate inbreeding in real populations? *Mol. Ecol.* 13, 3021–3031
21. Leffler, E.M. *et al.* (2012) Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biol.* 10, e1001388
22. Li, S. *et al.* (2014) Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biol.* 15, 557
23. Locke, D.P. *et al.* (2011) Comparative and demographic analysis of orang-utan genomes. *Nature* 469, 529–533
24. Xue, Y. *et al.* (2015) Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science* 348, 242–245
25. Prado-Martinez, J. *et al.* (2013) Great ape genetic diversity and population history. *Nature* 499, 471–475
26. Abascal, F. *et al.* (2016) Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biol.* 17, 251
27. Shafer, A.B.A. *et al.* (2015) Genomics and the challenging translation into conservation practice. *Trends Ecol. Evol.* 30, 78–87
28. Kardos, M. *et al.* (2016) Genomics advances the study of inbreeding depression in the wild. *Evol. Appl.* 9, 1205–1218
29. Ouborg, N. *et al.* (2010) Conservation genetics in transition to conservation genomics. *Trends Genet.* 26, 177–187
30. Allendorf *et al.* (2010) Genomics and the future of conservation genetics. *Nat. Rev. Genet.* 11, 697–709
31. Zhang, F. *et al.* (2009) Copy number variation in human health, disease, and evolution. *Annu. Rev. Genom. Hum. Genet.* 10, 451–481
32. Zhang, F. *et al.* (2009) Complex human chromosomal and genomic rearrangements. *Trends Genet.* 25, 298–307
33. Cho, Y.S. *et al.* (2013) The tiger genome and comparative analysis with lion and snow leopard genomes. *Nat. Commun.* 4, 2433
34. Rasmussen, M. *et al.* (2011) An aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334, 94–98
35. Romiguier, J. *et al.* (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515, 261–263
36. Bi, K. *et al.* (2013) Unlocking the vault: next-generation museum population genomics. *Mol. Ecol.* 22, 6018–6032
37. Palkopoulou, E. *et al.* (2015) Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. *Curr. Biol.* 25, 1395–1400
38. Der Sarkissian, C. *et al.* (2015) Evolutionary genomics and conservation of the endangered Przewalski's horse. *Curr. Biol.* 25, 2577–2583
39. da Fonseca, R.R. *et al.* (2015) The origin and evolution of maize in the Southwestern United States. *Nat. Plants* 1, 14003
40. Węcek, K. *et al.* (2017) Complex admixture preceded and followed the extinction of wisent in the wild. *Mol. Biol. Evol.* 34, 598–612
41. Wandeler *et al.* (2007) Back to the future: museum specimens in population genetics. *Trends Ecol. Evol.* 22, 634–642
42. Taylor, A. *et al.* (1994) Genetic variation of microsatellite loci in a bottlenecked species: the northern hairy-nosed wombat *Lasiorhinus krefftii*. *Mol. Ecol.* 3, 277–290
43. Mak, S. *et al.* (2017) Comparative performance of the BGISEQ-500 versus Illumina HiSeq2500 sequencing platforms for palaeogenomic sequencing. *Gigasience* 6, 1–13
44. Hung, C.M. *et al.* (2014) Drastic population fluctuations explain the rapid extinction of the passenger pigeon. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10636–10641
45. Staats, M. *et al.* (2013) Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. *PLoS One* 8, e69189
46. Carøe, C. *et al.* (2017) Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* Published online October 10, 2017. http://dx.doi.org/10.1111/2041-210X.12871
47. Mikheyev, A.S. *et al.* (2015) Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. *Nat. Commun.* 6, 7991
48. Orlando, L. *et al.* (2015) Reconstructing ancient genomes and epigenomes. *Nat. Rev. Genet.* 16, 395–408
49. Warinner, C. *et al.* (2014) Pathogens and host immunity in the ancient human oral cavity. *Nat. Genet.* 46, 336–344

50. Willerslev, E. *et al.* (2014) Fifty thousand years of arctic vegetation and megafaunal diet. *Nature* 506, 47–51
51. Prüfer, K. *et al.* (2010) Computational challenges in the analysis of ancient DNA. *Genome Biol.* 11, R47
52. Shapiro, B. and Hofreiter, M. (2014) A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. *Science* 343, 1236573
53. Pickrell, J.K. and Reich, D. (2014) Toward a new history and geography of human genes informed by ancient DNA. *Trends Genet.* 30, 377–389
54. Spötter, A. *et al.* (2016) Genome-wide association study of a varroa-specific defense behavior in honeybees (*Apis mellifera*). *J. Hered.* 107, 220–227
55. Lynch, M. (2008) Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. *Mol. Biol. Evol.* 25, 2409–2419
56. Briggs, A.W. *et al.* (2010) Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA. *Nucleic Acids Res.* 38, e87
57. Magi, A. *et al.* (2014) H3M2: detection of runs of homozygosity from whole-exome sequencing data. *Bioinformatics* 30, 2852–2859
58. McQuillan, R. *et al.* (2008) Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83, 359–372
59. Pemberton, T.J. *et al.* (2012) Genomic patterns of homozygosity in worldwide human populations. *Am. J. Hum. Genet.* 91, 275–292
60. de Valles-Ibáñez, G. *et al.* (2016) Genetic load of loss-of-function polymorphic variants in great apes. *Genome Biol. Evol.* 8, 871–877
61. Hedrick, P.W. and Garcia-Dorado, A. (2016) Understanding inbreeding depression, purging, and genetic rescue. *Trends Ecol. Evol.* 31, 940–952
62. Sudmant, P.H. *et al.* (2015) An integrated map of structural variation in 2,504 human genomes. *Nature* 526, 75–81
63. Rogers, R.L. and Slatkin, M. (2017) Excess of genomic defects in a woolly mammoth on Wrangel island. *PLoS Genet.* 13, e1006601
64. Huang, D.W. *et al.* (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13
65. Bottrill, M.C. *et al.* (2008) Is conservation triage just smart decision making? *Trends Ecol. Evol.* 23, 649–654
66. Fry, E. *et al.* (2017) Accumulation and functional architecture of deleterious genetic variants during the extinction of Wrangel island mammoths. *bioRxiv* Published online May 14, 2017. http://dx.doi.org/10.1101/137455
67. Schwartz *et al.* (2007) Genetic monitoring as a promising tool for conservation and management. *Trends Ecol. Evol.* 22, 25–33
68. Hansen *et al.* (2012) Monitoring adaptive genetic responses to environmental change. *Mol. Ecol.* 21, 1311–1329
69. Irizarry, K.J.L. *et al.* (2016) Integrating genomic data sets for knowledge discovery: an informed approach to management of captive endangered species. *Int. J. Genom. Proteom.* 2016, 2374610
70. Kosch, T.A. *et al.* (2017) Characterization of MHC class IA in the endangered southern corroboree frog. *Immunogenetics* 69, 165–174
71. Johnson, J.A. *et al.* (2016) Is there a future for genome-editing technologies in conservation? *Anim. Conserv.* 19, 97–101
72. Bouzat, J.L. *et al.* (1998) The ghost of genetic diversity past: historical DNA analysis of the greater prairie chicken. *Am. Nat.* 152, 1–6
73. Cahill, J.A. *et al.* (2015) Genomic evidence of geographically widespread effect of gene flow from polar bears into brown bears. *Mol. Ecol.* 24, 1205–1217
74. Frantz, L.A.F. *et al.* (2016) Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* 352, 1228–1231
75. Dobrynin, P. *et al.* (2015) Genomic legacy of the African cheetah, *Acinonyx jubatus*. *Genome Biol.* 16, 277
76. DePristo, M.A. *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498
77. Haubold, B. *et al.* (2010) mlRho – a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. *Mol. Ecol.* 19 (Suppl. 1), 277–284
78. Korneliussen, T.S. *et al.* (2014) ANGSD: analysis of next generation sequencing data. *BMC Bioinform.* 15, 356
79. Bryc, K. *et al.* (2013) A novel approach to estimating heterozygosity from low-coverage genome sequence. *Genetics* 195, 553–561
80. Link, V. *et al.* (2017) ATLAS: analysis tools for low-depth and ancient samples. *bioRxiv* Published online February 2, 2017. http://dx.doi.org/10.1101/105346
81. Purcell, S. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575
82. Yao, Y.C. (1988) Estimating the number of change-points via Schwarz' criterion. *Stat. Probab. Lett.* 6, 181–189
83. Narasimhan, V. *et al.* (2016) BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics* 32, 1749–1751
84. Wang, K. *et al.* (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164
85. McLaren, W. *et al.* (2016) The Ensembl variant effect predictor. *Genome Biol.* 17, 122
86. Kistler, L. *et al.* (2017) A new model for ancient DNA decay based on paleogenomic meta-analysis. *Nucleic Acids Res.* 45, 6310–6320
87. Briggs *et al.* (2007) Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* 104, 14616–14621
88. Gilbert, M.T.P. *et al.* (2007) Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. *Nucleic Acids Res.* 35, 1–10
89. Burrell, A.S. *et al.* (2015) The use of museum specimens with high-throughput DNA sequencers. *J. Hum. Evol.* 79, 35–44
90. Gilbert, M.T.P. *et al.* (2003) Characterization of genetic miscoding lesions caused by postmortem damage. *Am. J. Hum. Genet.* 72, 48–61
91. Hansen, A. *et al.* (2001) Statistical evidence for miscoding lesions in ancient DNA templates. *Mol. Biol. Evol.* 18, 262–265
92. Green, R.E. *et al.* (2010) A draft sequence of the Neandertal genome. *Science* 328, 710–722
93. Orlando, L. *et al.* (2013) Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499, 74–78
94. Pinhasi, R. *et al.* (2015) Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One* 10, e0129102
95. Meyer, M. and Kircher, M. (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* 2010, db.prot5448