BIOLOGICAL REVIEWS
Cambridge Philosophical Society
Biol. Rev. (2022), **97**, pp. 1511–1538.
doi: 10.1111/brv.12852

Global genetic diversity status and trends: towards a suite of Essential Biodiversity Variables (EBVs) for genetic composition

Sean Hoban[1‡,§], Frederick I. Archer[2§], Laura D. Bertola[3‡,§], Jason G. Bragg[4§], Martin F. Breed[5‡,§], Michael W. Bruford[6‡,§], Melinda A. Coleman[7§], Robert Ekblom[8‡,§], W. Chris Funk[9‡,§], Catherine E. Grueber[10‡,§], Brian K. Hand[11§], Rodolfo Jaffé[12§], Evelyn Jensen[13§], Jeremy S. Johnson[14§], Francine Kershaw[15§], Libby Liggins[16§], Anna J. MacDonald[17,†‡,§], Joachim Mergeay[18,19‡,§], Joshua M. Miller[20§], Frank Muller-Karger[21§], David O'Brien[22§], Ivan Paz-Vinas[23§], Kevin M. Potter[24§], Orly Razgour[25§], Cristiano Vernesi[26§] and Margaret E. Hunter[27‡,§*]

[1]*Center for Tree Science, The Morton Arboretum, 4100 Illinois Rt 53, Lisle, IL, 60532, USA*
[2]*Southwest Fisheries Science Center, NOAA/NMFS, 8901 La Jolla Shores Drive, La Jolla, CA, 92037, USA*
[3]*City College of New York, 160 Convent Avenue, New York, NY, 10031, USA*
[4]*Research Centre for Ecosystem Resilience, Australian Institute of Botanical Science, The Royal Botanic Garden Sydney, Mrs Macquaries Rd, Sydney, NSW, 2000, Australia*
[5]*College of Science and Engineering, Flinders University, University Drive, Bedford Park, SA, 5042, Australia*
[6]*School of Biosciences, Cardiff University, Cathays Park, Cardiff, CF10 3AX, Wales, UK*
[7]*Department of Primary Industries, New South Wales Fisheries, National Marine Science Centre, 2 Bay Drive, Coffs Harbour, NSW, 2450, Australia*
[8]*Wildlife Analysis Unit, Swedish Environmental Protection Agency, Blekholmsterrassen 36, Stockholm, SE-106 48, Sweden*
[9]*Department of Biology, Graduate Degree in Ecology, Colorado State University, 1878 Campus Delivery, Fort Collins, CO, 80523-1878, USA*
[10]*School of Life and Environmental Sciences, Faculty of Science, The University of Sydney, Carslaw Building, Sydney, NSW, 2006, Australia*
[11]*Flathead Lake Biological Station, 32125 Bio Station Ln, Polson, MT, 59860, USA*
[12]*Exponent, 15375 SE 30th Place, Suite 250, Bellevue, WA, 98007, USA*
[13]*School of Natural and Environmental Sciences, Newcastle University, Agriculture Building, Newcastle Upon Tyne, NE1 7RU, UK*
[14]*Department of Environmental Studies, Prescott College, 220 Grove Avenue, Prescott, AZ, 86303, USA*
[15]*Natural Resources Defense Council, 40 West 20th Street, New York, NY, 10011, USA*
[16]*School of Natural Sciences, Massey University, Ōtehā Rohe campus, Gate 4 Albany Highway, Auckland, Aotearoa, 0745, New Zealand*
[17]*Research School of Biology, The Australian National University, Acton, ACT, 2601, Australia*
[18]*Research Institute for Nature and Forest, Gaverstraat 4, 9500, Geraardsbergen, Belgium*
[19]*Aquatic Ecology, Evolution and Conservation, KULeuwen, Charles Deberiotstraat 32, box 2439, 3000, Leuven, Belgium*
[20]*Department of Biological Sciences, MacEwan University, 10700 104 Avenue, Edmonton, AB, T5J 4S2, Canada*
[21]*College of Marine Science, University of South Florida, 140 7th Avenue South, Saint Petersburg, Florida, 33701, USA*
[22]*NatureScot, Great Glen House, Leachkin Road, Inverness, IV3 8NW, UK*
[23]*Laboratoire Evolution et Diversité Biologique, Université de Toulouse, CNRS, IRD, UPS, UMR-5174 EDB, 118 route de Narbonne, Toulouse, 31062, France*

* Author for correspondence (Tel.: +001 352 264 3484; E-mail: mhunter@usgs.gov).
†Present address: Australian Antarctic Division, Department of Agriculture, Water and the Environment, 203 Channel Highway, Kingston, Tasmania, 7050, Australia
‡Conservation Genetics Specialist Group, International Union for Conservation of Nature (IUCN), 1196 Gland, Switzerland.
§Group on Earth Observation Biodiversity Observation Network.

[24]*Department of Forestry and Environmental Resources, North Carolina State University, 3041 Cornwallis Road, Research Triangle Park, NC, 27709, USA*
[25]*Biosciences, University of Exeter, Streatham Campus, Hatherly Laboratories, Prince of Wales Road, Exeter, EX4 4PS, UK*
[26]*Forest Ecology Unit, Research and Innovation Centre- Fondazione Edmund Mach, Via E. Mach, 1, San Michele all'Adige, 38010, (TN), Italy*
[27]*U.S. Geological Survey, Wetland and Aquatic Research Center, 7920 NW 71st Street, Gainesville, FL, 32653, USA*

ABSTRACT

Biodiversity underlies ecosystem resilience, ecosystem function, sustainable economies, and human well-being. Understanding how biodiversity sustains ecosystems under anthropogenic stressors and global environmental change will require new ways of deriving and applying biodiversity data. A major challenge is that biodiversity data and knowledge are scattered, biased, collected with numerous methods, and stored in inconsistent ways. The Group on Earth Observations Biodiversity Observation Network (GEO BON) has developed the Essential Biodiversity Variables (EBVs) as fundamental metrics to help aggregate, harmonize, and interpret biodiversity observation data from diverse sources. Mapping and analyzing EBVs can help to evaluate how aspects of biodiversity are distributed geographically and how they change over time. EBVs are also intended to serve as inputs and validation to forecast the status and trends of biodiversity, and to support policy and decision making. Here, we assess the feasibility of implementing Genetic Composition EBVs (Genetic EBVs), which are metrics of within-species genetic variation. We review and bring together numerous areas of the field of genetics and evaluate how each contributes to global and regional genetic biodiversity monitoring with respect to theory, sampling logistics, metadata, archiving, data aggregation, modeling, and technological advances. We propose four Genetic EBVs: (*i*) Genetic Diversity; (*ii*) Genetic Differentiation; (*iii*) Inbreeding; and (*iv*) Effective Population Size (N_e). We rank Genetic EBVs according to their relevance, sensitivity to change, generalizability, scalability, feasibility and data availability. We outline the workflow for generating genetic data underlying the Genetic EBVs, and review advances and needs in archiving genetic composition data and metadata. We discuss how Genetic EBVs can be operationalized by visualizing EBVs in space and time across species and by forecasting Genetic EBVs beyond current observations using various modeling approaches. Our review then explores challenges of aggregation, standardization, and costs of operationalizing the Genetic EBVs, as well as future directions and opportunities to maximize their uptake globally in research and policy. The collection, annotation, and availability of genetic data has made major advances in the past decade, each of which contributes to the practical and standardized framework for large-scale genetic observation reporting. Rapid advances in DNA sequencing technology present new opportunities, but also challenges for operationalizing Genetic EBVs for biodiversity monitoring regionally and globally. With these advances, genetic composition monitoring is starting to be integrated into global conservation policy, which can help support the foundation of all biodiversity and species' long-term persistence in the face of environmental change. We conclude with a summary of concrete steps for researchers and policy makers for advancing operationalization of Genetic EBVs. The technical and analytical foundations of Genetic EBVs are well developed, and conservation practitioners should anticipate their increasing application as efforts emerge to scale up genetic biodiversity monitoring regionally and globally.

Key words: biodiversity monitoring, environmental policy, indicators, metadata, interoperability, molecular ecology

CONTENTS

I INTRODUCTION

Governments, business, and society increasingly recognize that biodiversity contributes to ecosystem resilience and function, sustainable economies, and human well-being. Changes in biodiversity impact ecosystem services, often in ways that are hard to predict. To address this problem, government, non-government, and research organizations aim to quantify biodiversity trends, assess drivers of biodiversity change, measure responses of biological systems to interventions, and define policy options for preserving and restoring biodiversity. These activities require observational data collected across space and time on all levels of biological organization (genes, species and ecosystems). The Group on Earth Observations Biodiversity Observation Network (GEO BON) is tackling this challenge with the Essential Biodiversity Variable (EBV) concept, to help aggregate, harmonize, summarize, and interpret complex biodiversity observation data, especially from diverse data sources (Pereira *et al.*, 2013; Brummitt *et al.*, 2017; Navarro *et al.*, 2017; Schmeller *et al.*, 2017).

EBVs are a set of fundamental metrics that describe genetic, species, population, and ecosystem diversity, building on and harmonizing direct observations (e.g. species occurrences) to serve as a basis for indicators or inputs for modeling and forecasting the status and trends of biodiversity. EBVs track core components of biodiversity change, using highly standardized and scalable biodiversity observations. The concept of Genetic Composition EBVs goes beyond traditional methods and statistics and their practical use in biodiversity monitoring (Schwartz, Luikart & Waples, 2007; Mimura *et al.*, 2017; Leroy *et al.*, 2018); EBVs should provide the most essential information across space, time, and species. According to Pereira *et al.* (2013), a short list of EBVs that are standardized, recognized, and responsive to natural and anthropogenic drivers will allow for uptake of high-quality, comparable and scalable biodiversity data by an array of endpoint users. Similar to Essential Climate Variables (ECVs; Bojinski *et al.*, 2014), EBVs should be relevant, sensitive to change, generalizable, scalable, feasible, and have substantial accessible databases available (Schmeller *et al.*, 2018). EBVs must consider the fundamental characteristics of biological systems, and methods for collecting and analyzing observation data. EBVs should be applicable to all life forms, at multiple scales, and in all ecosystems. The implementation of EBVs allows for data to be synthesized into products such as indicators (e.g. the Species Habitat Index, https://geobon.org/ebvs/indicators/) for measuring progress on policy commitments, such as the Convention on Biological Diversity (CBD), Sustainable Development Goals, Global Strategy for Plant Conservation, and Intergovernmental Panel on Biodiversity and Ecosystem Services (IPBES; Schmeller *et al.*, 2015; Brummitt *et al.*, 2017; Navarro *et al.*, 2017).

Pereira *et al.* (2013) proposed six EBV 'classes' corresponding to ways that biodiversity can be categorized, observed, and managed: Genetic Composition, Species Populations, Species Traits, Community Composition, Ecosystem Structure, and Ecosystem Function. Each EBV class may comprise multiple EBVs that quantify different key attributes of that level of biodiversity; they should be complementary and comprehensive. For example, two EBVs are proposed in the Species Populations class (species distribution and species abundance; Jetz *et al.*, 2019), while five EBVs are proposed in the Species Traits class (phenology, morphology, reproduction, physiology, and movement; Kissling *et al.*, 2018). The EBV concept was also used for invasive species (Latombe *et al.*, 2017), and an ocean observation framework (Marine BON; Muller-Karger *et al.*, 2018). Schmeller *et al.* (2018) proposed six criteria for evaluating the suitability of potential EBVs. Progress is ongoing to achieve consensus on suitable EBVs for each class and to define workflows connecting data to EBVs – work that must be accomplished before EBVs can be used operationally (Schmeller *et al.*, 2017).

Genetic composition (within-species genetic variation) is a foundational component of biodiversity. It contributes to function and structure in all ecosystems and to resilience and/or productivity in agriculture, aquaculture and forestry systems (see Section I.3). Genetic composition can respond to environmental change rapidly and without changes in species abundance (Balkenhol *et al.*, 2016), and in some cases genetic data are cheaper and easier to gather than other methods of biodiversity monitoring (see Carroll *et al.*, 2018; Ferreira *et al.*, 2018; Parsons *et al.*, 2018; Wheat *et al.*, 2016). Nonetheless, observations of genetic composition have not been sufficiently mobilized for, nor integrated into, large-scale biodiversity conservation policy, sustainable development, and decision-making on resource management to date (Laikre *et al.*, 2010; Pierson *et al.*, 2016; Bruford *et al.*, 2017; Thomson *et al.*, 2021), even though it ultimately determines the capacity of populations to adapt to environmental changes and therefore underpins their long-term persistence (Reed & Frankham, 2003; see Section I.3). Although many countries have monitoring programs for species and ecosystem variables, few systematically monitor genetic variation; a recent report found that the status and trends of genetic diversity are rarely included in National Reports to the CBD (Hoban *et al.*, 2021). However, Scotland has committed to regular reporting on genetic diversity, with the first assessment published in 2020 (Hollingsworth *et al.*, 2020), while Sweden has committed to large-scale monitoring of genetic diversity (Posledovich, Ekblom & Laikre, 2021).

(1) Genetic data and infrastructure are ready for EBVs

While there are many reviews of genetic tools and statistics, and their practical use in biodiversity monitoring (Schwartz *et al.*, 2007; Mimura *et al.*, 2017; Leroy *et al.*, 2018), there has not yet been an effort to define clearly Genetic Composition EBVs (hereafter Genetic EBVs), summarize the practical framework for collecting and analyzing genetic observations *via* standardized procedures, and examine the challenges and opportunities that will ensure their usefulness for biodiversity monitoring. It is now timely to do so. Over the last 50 years, genetic data production and methods have advanced rapidly. Genetic data were first collected in the form of isozymes and allozymes from the 1970s to early 1990s, followed by DNA sequences, microsatellites and other technologies from the late 1980s to the present day, and large-scale next generation sequencing assessing genome-wide variation starting in the 2000s and increasing in use to the present. Detailed overviews of the temporal trends, applications and integration of the various types of molecular markers can be found in Schlötterer (2004), Ouborg *et al.* (2010), Putman & Carbone (2014), Garner *et al.* (2016), Garrido-Cardenas, Mesa-Valle & Manzano-Agugliaro (2018), and Leigh *et al.* (2021).

Genetic composition has been assessed for thousands of species globally with increasingly high-resolution population-level genetic data sets available at decreasing costs (Vranckx *et al.*, 2012; Romiguier *et al.*, 2014; Lawrence *et al.*, 2019). Although the changing technology and limited number of systematic monitoring programs does pose problems for spatiotemporal comparisons (see Sections III.1 on obtaining genetic data and V.2 on scale and standardization), recent studies are identifying gaps in genetic marker data availability (Miraldo *et al.*, 2016) and are jointly analyzing genetic summaries obtained across different species, studies and markers using standardization and normalization procedures [e.g. De Kort *et al.* (2021) for amplified fragment length polymorphism (AFLP)/microsatellites]. Further, hundreds of species now have had their genomes sequenced and assembled, which makes developing genetic markers and interpreting genetic data more efficient in these and related species (e.g. within the same genus). We acknowledge that, at present, genetic data cover only a tiny fraction of all species on Earth and that it will take substantial time and effort to fill the spatiotemporal and taxonomic gaps. However, these gaps and biases are true for other biodiversity data [e.g. detailed data for species' population sizes and trends is limited to few species (Boakes *et al.*, 2010; Fithian *et al.*, 2015; Hortal *et al.*, 2015; Meyer *et al.*, 2016; Jetz *et al.*, 2019)] and the rapid advances in genetic data accessibility mean that these gaps are swiftly being filled (e.g. the growth in the availability of whole-genome sequences illustrated in fig. 2 in Leigh *et al.*, 2021).

Numerous agencies routinely assess the genetic composition of high-priority species, including *via* genome sequencing. Advances in sequencing technology, data storage and analysis are opening new frontiers including DNA sequencing 'in the field' (Pomerantz *et al.*, 2018; Ovaskainen *et al.*, 2020; see Section V), as well as analyzing DNA from biological collections or paleospecimens to establish pre-human-impact baselines (Bi *et al.*, 2019). Non-invasive (e.g. hair and scat) and environmental (e.g. soil, water) sampling is also increasingly standardized, making genetic data more widely available. There are well-developed, standard data-storage formats and analysis software (often open source, such as R – see *Molecular Ecology Resources*, Volume 17, Issue 1: Population Genomics with R, and online Supporting Information Fig. S1 and Table S1). Methods for modeling genetic composition and extrapolation in space and time are also well advanced (see Section IV.2).

Alongside data production, advances are occurring in data stewardship. Owing to the research community's adherence to the 2011 Joint Data Archiving Policy, standardized data formats, and the existence of numerous large-scale data archives (Section III), genetic data are among the most accessible of all biodiversity data, although challenges remain, e.g. missing metadata, incomplete data files, and biases (Toczydlowski *et al.*, 2021). The ability to collate and aggregate genetic data sets and calculate genetic summary statistics for potential Genetic EBVs has been demonstrated (Section IV). Furthermore, consensus and uptake of metadata standards for biological collections are aligning genetic

data sets with the FAIR principles [Findable, Accessible, Interoperable, and Reproducible (https://www.go-fair.org/fair-principles/; Wilkinson *et al.*, 2016)], applying data schemes like the Minimum Information about any (x) Sequence (MIxS, https://gensc.org/mixs/) devised by the Genomics Standards Consortium, and establishing highly searchable metadatabases (Sections III and IV). Accessible and affordable genomic-level data, and digitization and curation of older genetic data (Sections IV and V), provide a pathway towards operational Genetic EBVs. In parallel, all the aforementioned advances have recently favored the emergence of a new field named 'macrogenetics', which aims at understanding patterns and drivers of genetic variation across large spatial, temporal, and/or taxonomic scales [hundreds to thousands of species (Blanchet, Prunier & De Kort, 2017; Leigh *et al.*, 2021)]. The fast-growing field of macrogenetics is paving the way for Genetic EBV deployment (see Section IV.1 for details).

(2) What is genetic composition and why is it important for biodiversity policy, conservation, and society?

We define the Genetic EBV class to encompass inherited components, such as variations in DNA sequence or epigenetic modifications (Rey *et al.*, 2020), that determine form and function and vary among individuals or populations of a species. This EBV class has relevance to global policy targets including CBD Aichi Target 13, 'minimize genetic erosion' and 'safeguard genetic diversity'; Sustainable Development Goal (SDG) 2.5, 'maintain genetic diversity', SDG 14 on marine resources and SDG 15, on ecosystems and biodiversity; and to Global Strategy for Plant Conservation Target 5, 'effective management [for] conserving genetic diversity' and Target 9, '70 per cent of genetic diversity … conserved'. *Genetic diversity* is mentioned in these policy mechanisms, and its multiple dimensions must be understood, monitored, and managed to achieve sustainable development, benefits, and conservation goals. Genetic EBVs recently have been linked with conservation decision-making in the publication 'Effective Biodiversity Indicators' by the UK Government (Henly & Wentworth, 2021) and the Food and Agriculture Organization (FAO) has incorporated the effective population size (N_e) as an indicator for forest genetic monitoring (Graudal *et al.*, 2020).

Genetic composition is a measure of within-species diversity, which helps species adapt, maintain fitness, avoid inbreeding depression, and underlies the breadth of species' niches and the diversity of biotic and abiotic interactions. Genetic diversity is the foundation of the three levels of biodiversity, supporting and complementing species and ecosystems diversity. Genetic diversity provides resilience against abrupt changes and allows species and ecosystems to adapt to changing environments, climates, and other challenges (including diseases). Ultimately, genetic variation allows species to develop into distinct and new lineages. Genetic diversity within and across populations supports ecosystem functions and contributes vital resources to society (e.g. seagrass and mangroves that serve both to protect coastal habitats and as nurseries for fish), and services such as carbon capture (Kettenring *et al.*, 2014; Hollingsworth *et al.*, 2020). Genetic diversity has also been used for millennia to provide thousands of domesticated or harvested species with new adaptations to climate, disease, soil types, food sources, altitude, etc., and to improve sustainability and reduce industrial inputs (e.g. natural defenses against pests). In short, genetic diversity underlies processes upon which all other biodiversity depends.

Genetic diversity data can be observed directly through the analysis of genetic markers (e.g. DNA sequences, microsatellite fragment data, single nucleotide polymorphisms (SNPs), methylation patterns, etc.) in individuals sampled through geographic space and time. Genetic observations are often summarized as allele frequencies at the population level (analogous to species abundances in an ecosystem; Hu, He & Hubbell, 2006), or through summary statistics measured at various hierarchies (individual, population/metapopulation, lineage, or species level; Fig. 1). Such hierarchies can be characterized as alpha diversity (population-level measures at one location), beta diversity (differentiation between locations in space or time), and gamma diversity (across the system of populations).

Note that molecular genetic tools such as DNA sequences are also used to gather observations for taxonomic diversity at and above the species level (e.g. phylogenetics and systematics). While genetic data within a species are applied to the Genetic EBV class, data above the species level are appropriate for the Species or Community EBV classes. Molecular genetic data may also be collected for genetic forensics and environmental DNA (eDNA) assessments. Targeted (e.g. species-specific) and metabarcoding eDNA tools can infer species presence using eDNA detected in water, soil, and air samples (Goodwin *et al.*, 2019; Balázs *et al.*, 2020) and are well suited to provide data for the Species Population EBV classes; but note, only within-species data are relevant for the Genetic EBV class.

(3) Aims and structure of this review

In this review we elucidate the major stages and gaps on the route towards enabling large-scale standardized observations of genetic variants. Section II defines Genetic EBVs, explains their calculation and meaning, and assesses how each Genetic EBV meets the criteria of Schmeller *et al.* (2018): relevance, sensitivity to change, generalizability, scalability, feasibility, and available data. Section III explains the workflow for transforming, aggregating, harmonizing and archiving genetic data underlying Genetic EBVs, including publishing data and metadata standards in interoperable formats for integration with large data sets. In Section IV, we discuss methods and models for interpolation and extrapolation, translating genetic data or proxies for genetic data into patterns of change across space and time, and looking towards EBV development for global application. In Section V, we

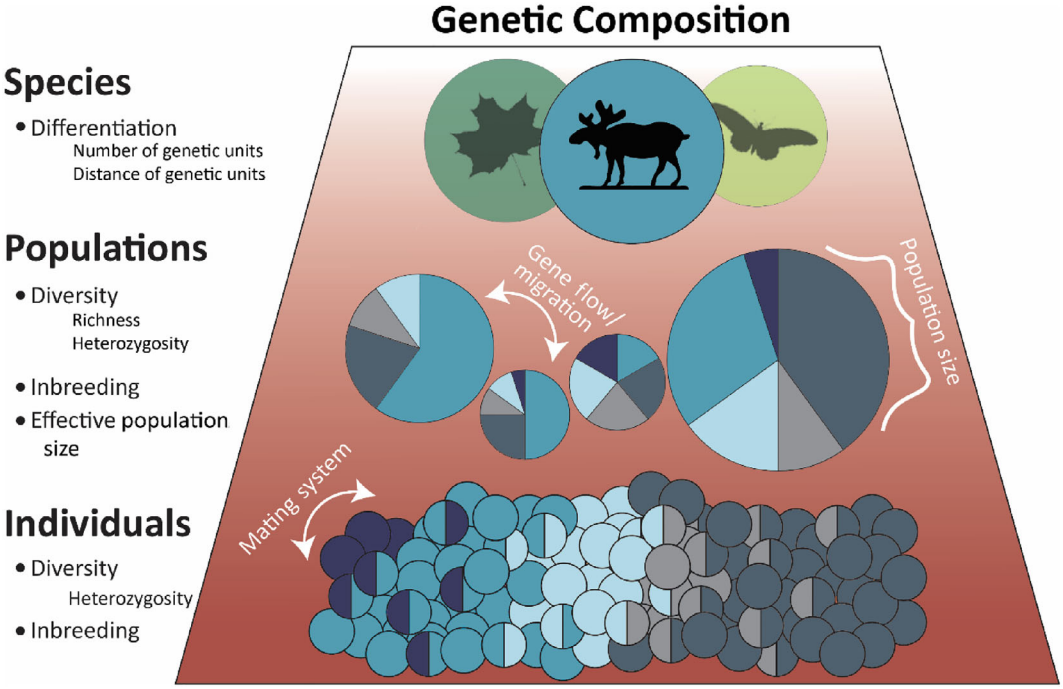


Fig. 1. The four Genetic Essential Biodiversity Variables (EBVs; bullet points) are indicated below each level of biological organization (Species, Populations, Individuals) for which they can be calculated. The species level corresponds to the combined genetic diversity for the species. The population level pie charts reflect the relative population sizes and the proportion of genotypes in each population (i.e. population genetic structure resulting from gene flow and migration). The smallest circles represent unique individuals with the colors depicting their genotypes.

discuss opportunities and challenges, and assimilation of Genetic EBVs into regional and global policy instruments. We conclude that the technical and analytical foundation of Genetic EBVs is extremely well developed, but additional steps are needed to scale Genetic EBVs for monitoring programs globally, establish best practices for informatics and modeling, and reconcile data sets with different resolutions.

II GENETIC EBVs

(1) Four proposed Genetic EBVs

Genetic EBVs have been proposed previously but were not thoroughly defined or examined in detail (Pereira *et al.*, 2013; Turak *et al.*, 2017; Kissling *et al.*, 2018). The four previously proposed EBVs were: allelic diversity, population genetic differentiation, co-ancestry, and breed/varietal diversity. Here we refine, revise and clarify these EBVs (Table 1). The first three are proposed to be renamed as: (*i*) Genetic Diversity, (*ii*) Genetic Differentiation, and (*iii*) Inbreeding. The first three EBVs reflect Wright's *F*-statistics (Wright, 1969, 1978) measuring the genetic variation among populations, within populations and within individuals, respectively. A new EBV, (*iv*) Effective Population Size (N_e), also developed by Wright, is added to measure the rate of change (particularly loss) of genetic diversity due to drift in the next generations (Luikart *et al.*, 2010; Tallmon *et al.*, 2010). As opposed to the current state of genetic diversity, N_e informs inferences on future genetic diversity changes and is highly suited for monitoring of conservation management actions. The previous EBV of breed/varietal diversity was integrated into the more comprehensive Genetic Differentiation EBV.

We examine how each EBV reflects the essential characteristics of genetic composition, genetic health, and genetic viability of populations and species. We explain the biological features that each represents, discuss how to interpret change, and review assumptions and cautions. The proposed list should be applicable for wild, semi-wild or managed, and domesticated species, and covers the forces that underpin genetic variation (Section II.1*a*). A key point is that although each EBV summarizes a different aspect of genetic composition, *all Genetic EBVs can typically be calculated from a single data set originating from DNA samples collected from individuals in a target population*. The EBVs we propose can be calculated using different genetic markers (e.g. whole-genome sequencing data, SNPs, DNA sequences, microsatellites).

(*a*) *What drives change in genetic composition?*

The field of population genetics has a well-established mathematical framework and predictive theory for how processes interact to shape genetic composition, which can enable a high degree of forecasting accuracy. The genetic variation of

Table 1. The four proposed Genetic Composition Essential Biodiversity Variables (EBVs), their definitions, possible values, interpretation, and a ranking from low (+) to high (+++) on various EBV criteria (rows)

Genetic EBV	Genetic Diversity		Genetic Differentiation		Inbreeding	Effective Population Size (N_e)
	Richness	Evenness	Number of genetic units	Differentiation between units		
Definition	Count of the number of alleles in a population	Expected proportion of heterozygotes in a population at equilibrium	Number of genetic lineages/ units within a species	Degree of genetic differentiation among populations or units	Degree of relatedness between pairs of individuals, mating among relatives, or identity by descent	Size of an ideal population that loses genetic variation at the same rate as the focal population
Examples	Allelic richness, nucleotide diversity (π)	H_e, H_O	Number of distinct evolutionarily significant units or management units	F_{ST}, G_{ST}, D_{est}	F_{IS}, F_{ROH}, F_H, t_m–t_s, t_m	LD-based N_e; N_e from temporal method
Range of values	0 to 1	0 to 1	≥1	0 to 1	-1 to 1	Usually smaller than census population size
Level of organization	Population	Individual/ population	Across species range or specific region	Across species range or specific region	Individual (pairs or families)/population	Population
Interpretation of a change identified across loci	Decline indicates loss of adaptive capacity	Decline indicates loss of adaptive capacity or increased levels of inbreeding	Loss of independent evolutionary or demographic units (probable loss of adaptations and long-term persistence)	Increased differentiation indicates more fragmentation (reduced gene flow) between units; decreased differentiation indicates homogenization	Increased inbreeding increases the chance of expression of deleterious alleles, reducing fitness	Lower N_e increases genetic drift, inbreeding, and future loss of genetic variation
Relevance to species status	+++	++	+++	++	+++	+++
Sensitivity to change over time	+++	+	+++	++ (depends on life history and N_e)	+++ (depends on method)	++
Spatial scalability*	+++	+	+++	++	+	+
Feasibility	+++	+++	+++	+++	++ (depends on measure – F_{ROH} currently less feasible)	+/++ (depends on method – single sample is more feasible
Data available	+++	+++	+++	+++	+	+

H_e = expected heterozygosity under Hardy–Weinberg assumptions, and H_O = observed heterozygosity, yielding the probability of randomly drawing two different alleles from the population. F_{ST} = genetic differentiation (fixation index) measured as the proportion of the total genetic variance contained in a subpopulation relative to the total genetic variance. G_{ST} = genetic differentiation of subpopulations relative to the total population, a generalized index of F_{ST} for loci with two or more alleles. D_{est} = unbiased genetic differentiation among subpopulations. F_{IS} = genetic differentiation of individuals relative to the subpopulation, reflecting deviations from random mating within subpopulations (inbreeding). F_{ROH} = inbreeding by autozygosity estimated from runs of homozygosity, i.e. the proportion of the autosomal genome above a specified length. F_H = the increase in individual homozygosity relative to mean Hardy–Weinberg expected homozygosity. t_m = the multilocus outcrossing rate, which is 1 – the self-fertilization rate. t_s = the degree of outcrossing among unrelated individuals only. t_m–t_s = the estimated rate of biparental inbreeding. LD, linkage disequilibrium.
*See Appendix SI for discussion of alternative spatial scalability ranking. Note that each EBV can be represented by numerous statistics which have individual calculations (Leroy *et al.*, 2018; see Appendix S1).

Biological Reviews **97** (2022) 1511–1538 © 2022 The Authors. *Biological Reviews* published by John Wiley & Sons Ltd on behalf of Cambridge Philosophical Society. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

species and their populations is entirely shaped by four processes: (*i*) *Mutation* is the random creation of new genetic variants, increasing genetic diversity. (*ii*) *Genetic drift* is the fluctuation in a population's allele frequencies due to the random sampling of alleles from generation to generation (e.g. not due to selection). In small populations, these chance statistical effects can lead to rapid change and typically decrease genetic diversity (e.g. 'genetic erosion') and increase among-population genetic differentiation over time (although note that genetic drift can reduce the efficacy of stabilizing selection, which might increase diversity). (*iii*) *Gene flow* is the exchange of genetic variation among populations *via* gametes (e.g. pollen, sperm) or whole organisms that mate and produce offspring, and can drive changes in local levels (including local increases) of genetic variation. (*iv*) *Selection* is a process influencing the differential survival and reproduction of genetic variants across space and time, and can influence the genetic variation within populations and the differentiation between them, depending on the type of selection. Selection is related to specific genomic sites (the genes influencing certain phenotypes, and linked regions), while drift, mutation, and migration can affect diversity at sites across the genome. The influence of these four processes is mediated by N_e, which is roughly related to the number of breeding adults in a genetically ideal population. High N_e results in more effective selection, more mutations, and lower genetic drift (Charlesworth, 2009). Decades of genetic measurements of these four processes have identified and quantified major detrimental anthropogenic drivers (e.g. harvesting, selective breeding, habitat change, pollution, reduced population size, isolation, and habitat fragmentation), as well as beneficial anthropogenic drivers such as genetic rescue and migration corridors (Pacioni, Wayne & Page, 2019; Almeida-Rocha *et al.*, 2020). Genetic composition is also influenced by life history traits that render some species more sensitive to genetic erosion, or less able to respond to environmental changes through genetic adaptations (e.g. long generation time, small population sizes, inbreeding, self or clonally propagating, restricted dispersal ability) (Romiguier *et al.*, 2014; Schmeller *et al.*, 2018).

While, the genetic EBVs are not presented in order of importance, we begin with Genetic Diversity within a population because this is the unit where natural selection takes place. Nonetheless we emphasize that defining populations (e.g. assessing genetic differentiation or population structure) is usually the first step in evaluating Genetic EBVs. Genetic and other data (environmental, geographic location) can be used to help define populations, a topic covered extensively in the literature (Fenderson, Kovach & Llamas, 2020). Once independent population units are defined, the other EBVs are calculated within each unit. EBVs would typically be calculated on genome-wide or neutral-marker data sets, although see the following discussion on adaptive diversity.

(*b*) *A note on adaptive diversity*

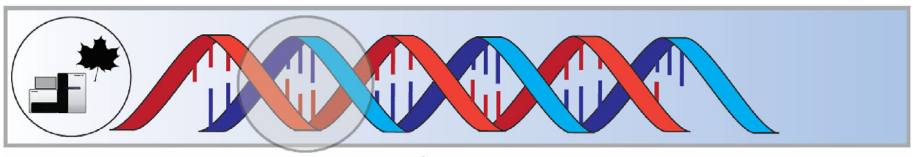
Genetic loci are often categorized as adaptive (e.g. 'non-neutral'), or 'neutral'. Adaptive genetic differentiation, the outcome of natural selection, may confer fitness advantages under local conditions. Genetic composition will differ when measured using adaptive compared to neutral genetic markers (Ralls *et al.*, 2018), which can impact recommended management actions (e.g. Van Oppen *et al.*, 2017). When possible, adaptive loci should be identified and analyzed separately, as comparing population genetic structure assessed using neutral and adaptive markers can provide alternative conservation options (Waples & Lindley, 2018). Identifying adaptive loci under selection is challenging (Hoban *et al.*, 2016), but methods are advancing to account better for loci under selection (Hohenlohe, Funk & Rajora, 2021). Various correlation methods can be used to identify putative adaptive loci, often through interactions with simple environmental data (Rellstab *et al.*, 2015; Ahrens *et al.*, 2018). Alternatively, loci with a particular, known function (identified with a reference genome and/or field and laboratory experiments) can be analyzed, although such data are available for fewer species. Nonetheless, these loci (assuming sufficiently large SNP datasets, good ecological knowledge and a sequenced, annotated genome) are useful to map local adaptations, model adaptive responses, and determine the number and degree of differentiation among 'population units' – the required first step for evaluating Genetic EBVs.

(2) Genetic diversity EBVs: richness and evenness

'Genetic diversity' describes the level of genetic variability within a population that enables species to adapt and persist. Previously, Pereira *et al.* (2013) referred only to allelic richness, but genetic diversity has two key components: richness and evenness; both should be recorded to capture genetic diversity fully. Both are easy metrics to assess and explain to practitioners (Fig. 2). Richness refers to the count of genetic variants in a population, corrected for sample size (e.g. an expectation for a given sample size). Several statistics fall under this category including number of haplotypes (usually for organelle DNA), number of distinct genetic sequences observed in a sample, and allelic richness [for microsatellites and SNPs, the mean number of genetic variants (alleles) per defined genomic position (locus)]. The number and difference among genetic states are the units that ultimately determine form and function and are the units on which natural selection can act. Genetic richness can only be increased *via* mutations (or, locally, by gene flow).

Evenness of genetic diversity is often represented by expected heterozygosity (H_e). It is probably the most widely reported statistic in genetic studies, irrespective of the marker type. It expresses the probability that two randomly drawn alleles from a sample are different (it is identical to the Gini–Simpson diversity index from ecological literature). Similar to richness, evenness will help a population respond rapidly to environmental change through adaptation. Both richness and evenness have been recorded in thousands of genetic studies (see Section IV).

Richness and evenness are correlated in populations at mutation-drift equilibrium (Nei, Maruyama & Chakraborty,



Individual	Sequence	Genotype L1	Genotype L2	Genotype L3	Genotype L4	Minor allele count L1	Minor allele count L2	Minor allele count L3	Minor allele count L4	By sample H_o
S1	GAATCTGAACCA GGATCTGATCCA	AG	CC	AT	CC	1	0	1	0	0.5
S2	GAATCTGATCAA GAATGTGATCCA	AA	CG	TT	AC	0	1	0	1	0.5
S3	GGATCTGAACCA GGATCTGATCCA	GG	CC	AT	CC	2	0	1	0	0.25
S4	GAATCTGATCCA GAATCTGATCCA	AA	CC	TT	CC	0	0	0	0	0
By locus H_o		0.25	0.25	0.5	0.25					

Fig. 2. A sample of four diploid individuals from a population, with various representations of genetic composition data structures. The workflow process includes genetic sequencing, aligning the sequences from each individual, and polymorphic loci identification. The data from the polymorphic sites (single nucleotide polymorphisms; SNPs) in the sequence can be summarized as a matrix of genotypes for each locus (L1–L4). When these loci are bi-allelic SNPs, the data can be summarized as the Minor Allele Count – the number of occurrences of the least frequent allele at that locus, a convenient summary format for certain statistics and models. A Genetic Diversity EBV for evenness, such as observed heterozygosity (H_o), can also be summarized for each individual (rows), or by locus (columns), as illustrated for the SNP matrix. Note that some measures of diversity include invariant sites which are calculated from sequence alignments, not a matrix of SNP genotypes.

1975; Fischer *et al.*, 2017), but these metrics respond differently to perturbations. Allelic richness measures respond faster to decreases in population size (similar to loss of rare species) while genetic evenness (heterozygosity) may respond faster to population mixing. Such different responses can be used to make inferences on past demography (population expansions or bottlenecks; Nei *et al.*, 1975). For both measures, a decrease is generally interpreted as a loss of evolutionary potential – the ability of a population or species to adapt to rapid environmental changes such as climate change. Importantly, genetic measures of richness and evenness are not directly comparable when calculated using different genetic markers (although standardization procedures can make them comparable; see De Kort *et al.*, 2021), so interpretation should be directed by the type and number of genetic markers and the number of populations/individuals analyzed (see Section V.2 on data set parameter selection). Alternatively, an integrated framework for measuring diversity, such as Hill numbers, could be explored for unifying different measures of Genetic EBVs, and to investigate variation at different EBV classes across multiple scales (Sherwin *et al.*, 2017; Gaggiotti *et al.*, 2018).

(3) Genetic differentiation EBVs: number of units and their connectivity

Genetic differentiation is defined as a divergence in genetic composition (specifically the frequencies of alleles) between multiple populations of the same species, caused by natural selection, genetic drift, and restricted gene flow (Section II.1*a*). This is a measure of beta diversity. The Genetic Differentiation EBV includes both the number of distinct populations (or evolutionary lineages) and the levels of differentiation between them; both are needed to capture and interpret 'differentiation' fully. Genetic differentiation is sometimes also referred to as population structure, and can be organized hierarchically (populations, regions, management units, genetic clusters, etc.). The previously proposed EBV of cultivated 'varietal or breed diversity' in reference to domesticated species could fall under the term 'genetic differentiation', which can apply to all species. The term 'genetic connectivity' typically connotes gene flow (see Section II.1*a*), which usually reduces genetic differentiation.

Genetic differentiation can be assessed *via* clustering methods [e.g. STRUCTURE (Pritchard, Stephens & Donnelly, 2000); discriminant analysis of principal components (DAPC) (Jombart, Devillard & Balloux, 2010); see Table S1] that identify the number and genetic composition of groups, or methods which compute summary statistics of the degree of differentiation among *a priori* defined groups, usually on a scale of 0 to 1 (see Table 1). It is important to note that measures of genetic differentiation are usually based on *a priori*, or at times *a posteriori*, stratification of samples representing prospective units. As such, these units can only represent hypotheses of genetic structuring, which may not correspond to all underlying patterns of differentiation. Thus, this EBV is measured under the assumption that

population units have been appropriately delineated and provides an essential context within which the other three Genetic EBVs are assessed.

Loss of population units is clearly detrimental from a conservation perspective. An increase in differentiation is also typically detrimental from a manager's perspective, as it is caused by a decrease in population size and/or a decrease in connectivity (e.g. habitat fragmentation or isolation). However, a decrease in differentiation may sometimes be detrimental, for example when isolated populations adapted to contrasting environments are homogenized, or when a managed species hybridizes with wild populations. In other cases, a decrease in differentiation may be the beneficial outcome of targeted genetic rescue (Whiteley *et al.*, 2015), whereby isolated populations are connected through corridors or translocations. Genetic differentiation is relatively easy to calculate, visualize and explain to conservation managers and policy makers. Unfortunately, many measures of genetic differentiation show slow responses to perturbation in some situations (tens of generations; e.g. Landguth *et al.*, 2010), have undesirable statistical properties (Jost *et al.*, 2018), and are difficult to apply in species with continuous distributions or asymmetric gene flow. Alternative individual-based measures can be used as proxies for gene flow (Shirk, Landguth & Cushman, 2017). Moreover, current differentiation levels may already be a result of habitat fragmentation or other human influences.

Highly genetically differentiated populations, or groups of populations, such as adaptively or morphologically distinct varieties or sets of populations, are often called Evolutionarily Significant Units (ESUs). These have been segregated for hundreds of generations and are usually defined by a set of genetically based traits or adaptations (Funk *et al.*, 2012). Such independent genetic units may comprise a breed or cultivar in domesticated species, or an ecotype, subspecies, variety, or stock in non-domesticated species. The number of such units is essential, as loss of ESUs equates to loss of a distinct product of evolutionary history such as unique traits and co-adapted sets of traits. In practice, the units are delineated by significant divergence at multiple loci indicative of independent evolutionary trajectories plus some ecological or trait differences (Fraser & Bernatchez, 2001; de Queiroz, 2007). There is increasing effort to identify and conserve ESUs at national and international scales, promoting evolutionary processes (e.g. Gene Conservation Units in European forestry; Koskela *et al.*, 2013).

(4) Inbreeding EBV

Measurements of inbreeding can be used to determine the level of diversity *within individuals*. Mating between related individuals constitutes inbreeding. Highly inbred individuals often have lower fitness (termed 'inbreeding depression'), which can compromise population growth and persistence by reducing demographic rates (e.g. decreased fecundity). Inbreeding metrics indicate near-term risks, and can signal that management action, such as translocating individuals, is urgently needed. When populations are large and/or there is sufficient gene flow within and between populations, inbreeding rates tend to be reduced. This EBV was previously called 'co-ancestry' (Pereira *et al.*, 2013) but 'inbreeding' is more reflective of the process and is an accessible concept for practitioners. Inbreeding is often a consequence of small population size, increasing rapidly when N_e is less than approximately 50 and in fragmented populations where gene flow dynamics are impacted (e.g. plant pollination; Breed *et al.*, 2015). However, some species can have naturally high inbreeding levels due to their mating system without any apparent fitness loss and some benefit (e.g. reproductive assurance in plants; Winn *et al.*, 2011).

Inbreeding can be estimated using various methods depending on the timescales and type of data available: path analysis on a pedigree; the degree of homozygosity using whole-genome data (e.g. runs of homozygosity); direct observations of inbreeding *via* the genotyping of progeny arrays; or the inbreeding coefficient, F_{IS}, which provides evidence of non-random mating according to Hardy–Weinberg equilibrium expectations [i.e. excess in observed homozygous genotypes (Kardos, Luikart & Allendorf, 2015; Kardos *et al.*, 2018)]. In addition to these measures of contemporary inbreeding, one can measure risks of inbreeding. Relatedness, kinship and co-ancestry metrics can identify related individuals in a population and quantify the likelihood of inbreeding before it occurs. Estimates of inbreeding have one major advantage over diversity and differentiation: estimates made with a given method are easy to compare across studies, as they should not be dependent on the type of DNA markers used, as long as the data set has sufficient statistical power.

(5) Contemporary effective population size EBV

Effective population size (N_e) is an established fundamental parameter for genetic biodiversity monitoring. In a population, N_e quantifies the amount of genetic change that is occurring *and* will occur in the future as a result of genetic drift. N_e determines the equilibrium level of genetic diversity that a population can maintain and thus the long-term genetic diversity in a population (Leffler *et al.*, 2012; Ellegren & Galtier, 2016). N_e is the most important parameter for assessing the ability of populations to adapt to rapid environmental change and the likelihood of inbreeding depression and other reactions to perturbations (Hoban, 2014; Tenesa *et al.*, 2007). For example, depending on the scenario, other less-sensitive or well-suited metrics, such as H_e, may have minimal decline between sampling timepoints, while N_e may decline more rapidly and indicate that rapid loss in genetic variation is predicted in the near future.

N_e is defined as the number of individuals in a theoretically ideal population (e.g. most adults randomly mate and contribute a Poisson-distributed number of offspring every generation) that will experience the same magnitude of genetic diversity loss (due to random genetic drift) as the real population. N_e determines the influence of genetic drift based on the

amount of genetic diversity – when N_e is small, genetic diversity is lost from a population faster over time and the random fluctuations in allele frequency caused by genetic drift can overrule the effect of natural selection (Charlesworth, 2009). N_e is widely adopted in practical management due to its relationship to both short-term inbreeding and long-term evolutionary potential (Jamieson & Allendorf, 2012).

N_e can be estimated across different spatial and temporal scales (Ryman, Laikre & Hössjer, 2019). Historical N_e is calculated in terms of coalescence, which translates to an average allelic sampling variance across generations due to past genetic drift (and/or sampling variance) (Sjödin *et al.*, 2005). This provides a retrospective view of N_e equally across the entire genealogy, and limits the suitability of historic N_e for the assessment of change as an EBV. Historic N_e could be used to compare the average N_e for a genealogy across different populations or to characterize the diversity across multiple species in a region; however, these are less informative for monitoring diversity, especially as related to change influenced by environmental drivers (Nadachowska-Brzyska, Konczal & Babik, 2022).

Alternatively, in natural resource management situations, the more commonly applied N_e, and the primary focus of our EBV, is the *contemporary* N_e (Nadachowska-Brzyska *et al.*, 2022). In conservation situations, where recent demographic changes are the rule rather than the exception, an EBV addressing contemporary N_e can inform inferences on future genetic diversity changes. Using contemporary N_e, a prospective view of a population's genetic diversity in *n* generations is provided.

Two categories exist for contemporary N_e (hereafter N_e) estimators using genetic data – those requiring a single-time-point data set [e.g. linkage disequilibrium (LD), kinship-based methods], and those requiring two time points to be sampled, defined as temporal N_e estimates (Do *et al.*, 2014). The two categories can give concordant estimates, and both are influenced by marker and sample sizes (Ackerman *et al.*, 2017; Hoban *et al.*, 2020; Luikart *et al.*, 2021; Waples, 2021). Temporal methods progressively have been replaced by single-time-point methods since the late 2000s (Waples, 2021). In particular, linkage disequilibrium-based N_e estimates can be generated from a single time point and are more robust to violations in model assumptions and more easily comparable across studies (Waples & England, 2011), but are sensitive to population structure (Neel *et al.*, 2013). The different contemporary N_e estimations are complementary to one another and the use of several estimators can help to reduce uncertainty. There are also multiple kinds of N_e that quantify other parameters and apply to different timescales (e.g. inbreeding, variance, etc.). See Waples (2016) and Ryman *et al.* (2019) for in-depth reviews of the definitions and interpretation of each measure of N_e. When populations are stable and at mutation–drift equilibrium, the historical N_e is the same as the contemporary N_e and the latter reflects the allelic sampling variance of the previous generation(s). In all cases, N_e should be estimated within demographic units (metapopulations or populations).

Effective population size is clearly linked to the census population size (N_c) and the variance in reproductive success among individuals: the larger this variance, the smaller resultant N_e for a given N_c (Kimura & Crow, 1963). The effective size N_e is often between 1/2 and 1/10 of the N_c, but in some cases it is much smaller. By applying the average ratio for vertebrates of approximately $N_e/N_c = 0.1$, N_c can be inferred from genetically informed effective population size (see Hoban *et al.*, 2020).

N_e estimation is sensitive to sample size variance and mixed-age populations and requires careful application and interpretation. However, N_e is one of the best-studied and easily explained metrics for applying biodiversity loss thresholds: N_e below approximately 50 will lead to rapid increases in inbreeding, loss of fitness, and genetic composition change; while N_e below approximately 500 will result in loss of ability to adapt *via* natural selection (Frankham, 2005; Jamieson & Allendorf, 2012; but see also Frankham, Bradshaw & Brook, 2014 for arguments for doubling these thresholds to 100 and 1000, respectively). N_e is useful to assess conservation management needs and to monitor success of recovery programmes, including through indicator metrics and forecasting (Hoban *et al.*, 2020). From an applied perspective, N_e is a measure of population size, which in many cases is easier (and more cost effective) to determine using genetic tools than through traditional survey efforts.

(6) Why four Genetic EBVs?

The four proposed Genetic EBVs cover the components of Wright's genetic variation (Fig. 3; Jost, 2006; Daly, Baetens & De Baets, 2018) and together provide a comprehensive description of the impacts of environmental change on genetic composition. As with other classes of EBVs, various Genetic EBVs will sometimes all be affected similarly by a driver (environmental or human factors), while other drivers will lead to idiosyncratic changes. This provides some ability to distinguish among drivers, and to elucidate the underlying mechanistic process, by calculating and comparing multiple EBVs. For example, decreases in census size will likely lead to increases in inbreeding (and genetic drift), as well as associated loss in allelic richness and genetic evenness (heterozygosity). Changes in gene flow will lead to changes in genetic differentiation, average relatedness, allelic richness, and heterozygosity, but each on different timescales (Landguth *et al.*, 2010). Inbreeding and allelic richness typically change relatively quickly after the onset of disturbances, compared to genetic evenness (heterozygosity) and differentiation, which respond more slowly (Keyghobadi *et al.*, 2005; Lowe *et al.*, 2005). Note that interpretation of N_e and genetic differentiation can be sensitive to genetic data set sizes and also to levels of breeding between close relatives (i.e. biparental inbreeding) or selfing (primarily in plants). Note also that hybridization rate was considered as a potential EBV but was determined not to fit our criteria (see Appendix S2).

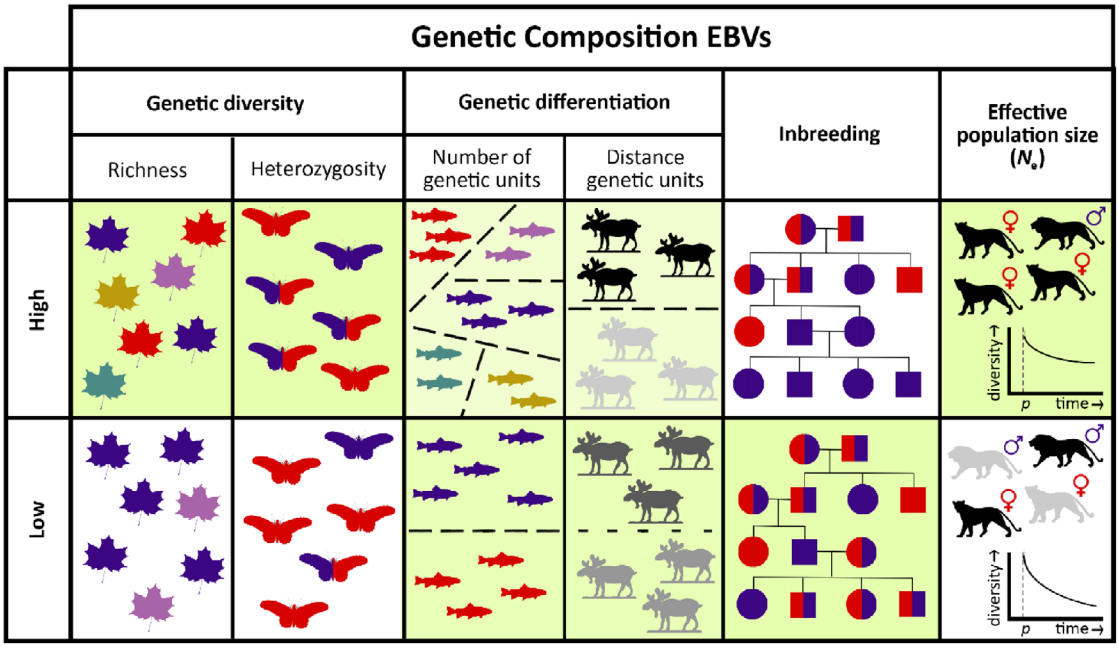


Fig. 3. The four Genetic Composition Essential Biodiversity Variables (EBVs). Green background shading indicates the preferred genetic state (high or low levels) in many conservation/management situations. The preferred state for genetic differentiation is context dependent, represented by a lighter shade of green (see text). Distance of genetic units illustrates high genetic distance (black *versus* white), and low genetic distance (dark gray *versus* light gray). The contemporary effective population size (N_e) is represented with black (breeding) and gray (non-breeding) individuals and the graphs denote projections after the present time (p) of the future losses of genetic diversity.

(7) Important notes on Genetic EBVs as biodiversity metrics

Here we highlight several relevant characteristics of Genetic EBVs. Due to the continuously inherited nature of DNA, Genetic EBVs are useful for assessing populations today and also their history at timescales from recent generations to millions of years ago. Geneticists can use DNA from contemporary samples to infer past changes in migration rates or the rate of change in N_e through models (see Section IV.2). For example, analysis of DNA in the rarest bat in the UK, *Plecotus austriacus*, showed a recent 30-fold decline in N_e at a median of 200 years ago, corresponding to changes in farming practices and disappearance of unimproved grasslands (Razgour *et al.*, 2013). Also, because DNA can be obtained from preserved remains up to tens of thousands of years old, the past state of genetic composition can be directly assessed. For example, the genome sequence of a single member of the Wrangel Island population of wooly mammoth (*Mammuthus primigenius*), the last remaining population before the species went extinct, showed 20% less heterozygosity and 28 times more runs of homozygosity (a metric of inbreeding) than a more ancient, healthy population (Palkopoulou *et al.*, 2015). This is not a unique characteristic of genetic composition data, as data for other EBV classes can be obtained from past relics (e.g. population sizes from fossil abundance, community composition from pollen in sediment). However, the amount of information in genetic data from even one individual can provide high-resolution, reliable inferences on past states and changes that are probably unmatched by other historical evidence.

Another unique aspect of genetic data is that a single genetic or genomic data set can be used for calculating many (or all) Genetic EBVs (assuming sufficient sampling of individuals). This is similar to Species EBVs in which individual observations underlie abundance and extent, but dissimilar to Trait EBVs which require distinct observations using various instruments to cover a range of EBVs. For example, Henry *et al.* (2009) collected DNA from scat of the Amur tiger (*Panthera tigris altaica*) population and genotyped 95 individuals at 12 microsatellite markers. This data set revealed strong genetic differentiation between two isolated regions, and very low allelic richness, heterozygosity (evenness) and N_e relative to other tiger populations [see Silva *et al.* (2020) for a plant example using SNPs]. Small sampling efforts, such as genomic sequences from fewer than 10 individuals, can assess the present state of EBVs and their change over time (Díez-del-Molino *et al.*, 2018).

Lastly, most Genetic EBVs are interpreted in a context of change. As with other EBV classes, they typically require samples across multiple points in space and/or time (with representative sampling and comparable data, e.g. similar sampling and analyses). However, some Genetic EBVs also have relevance based on their 'state value' *in addition to* 'change'. N_e and inbreeding can be calculated from single-time-point samples and have thresholds in their state value that signal vulnerability (Section II.5). For other EBVs, the

state value has meaning if comparing species with similar life histories but different exposure to drivers such as harvest (see discussion of third category in Section IV.1). Thus, to some extent, Genetic EBVs can be used when only single time points are available, although temporal monitoring is needed to determine the rate of change and degree of change from baselines.

III OBTAINING AND ARCHIVING GENETIC COMPOSITION DATA AND METADATA

(1) Obtaining genetic data

The steps in generating Genetic EBVs include collection of biological samples, laboratory work, computational processing of raw genetic data, analysis/calculation, publishing, archiving, modeling and/or synthesis (Fig. 4). Collecting genetic data begins with a biological sample – some remnant of a living or preserved individual, including 'invasive' sampling (direct handling of the organism to obtain a sample) or 'non-invasive' sampling (from hair, faeces, shed skin, etc.) (Fig. 5). Samples may be obtained from extant populations, biological collections, or sub-fossils (ancient DNA). Molecular biology techniques are used to analyze DNA sequences extracted from the sample. These data can be presented as sequences of DNA nucleotides, the allele 'states' at a locus for each individual, and/or as allele frequencies in populations, from which genetic statistics are derived (Fig. 2). DNA storage, extraction, sequencing and quality control are discussed briefly in Appendix S3. Data sets are often considered 'genetic' (tens of genetic markers, usually without a reference genome – as for most legacy data sets) or 'genomic' (thousands to millions of markers, often with a reference genome – as for many current data sets). Resolution and precision of EBV values are higher for the latter. Genetic composition can be reliably described for both data types, although reliable estimates of relatedness and inbreeding are best measured using genomic data sets. The growing size and complexity of genomic data sets creates the need for a corresponding growth in computer storage and power (i.e. supercomputers) and statistical programs to allow for

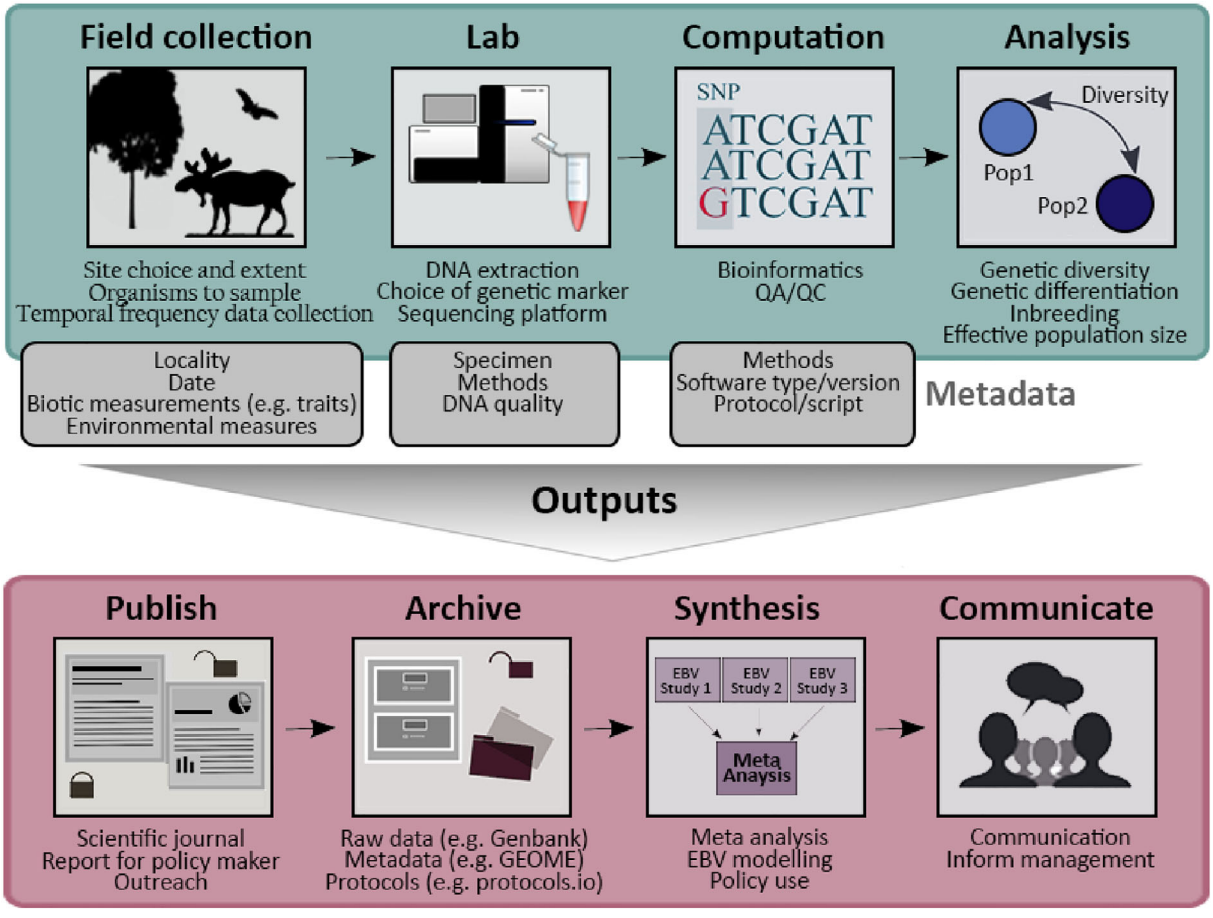


Fig. 4. Steps in generating Genetic Composition Essential Biodiversity Variables (EBVs) include field (or archive) collection of DNA, laboratory work, computational processing of raw data, analysis/calculation, publishing, archiving, modeling and/or synthesis and communication to inform management decisions. GEOME, Genomic Observatories MetaDatabase; Pop, population; QA/QC, quality assurance/quality control.

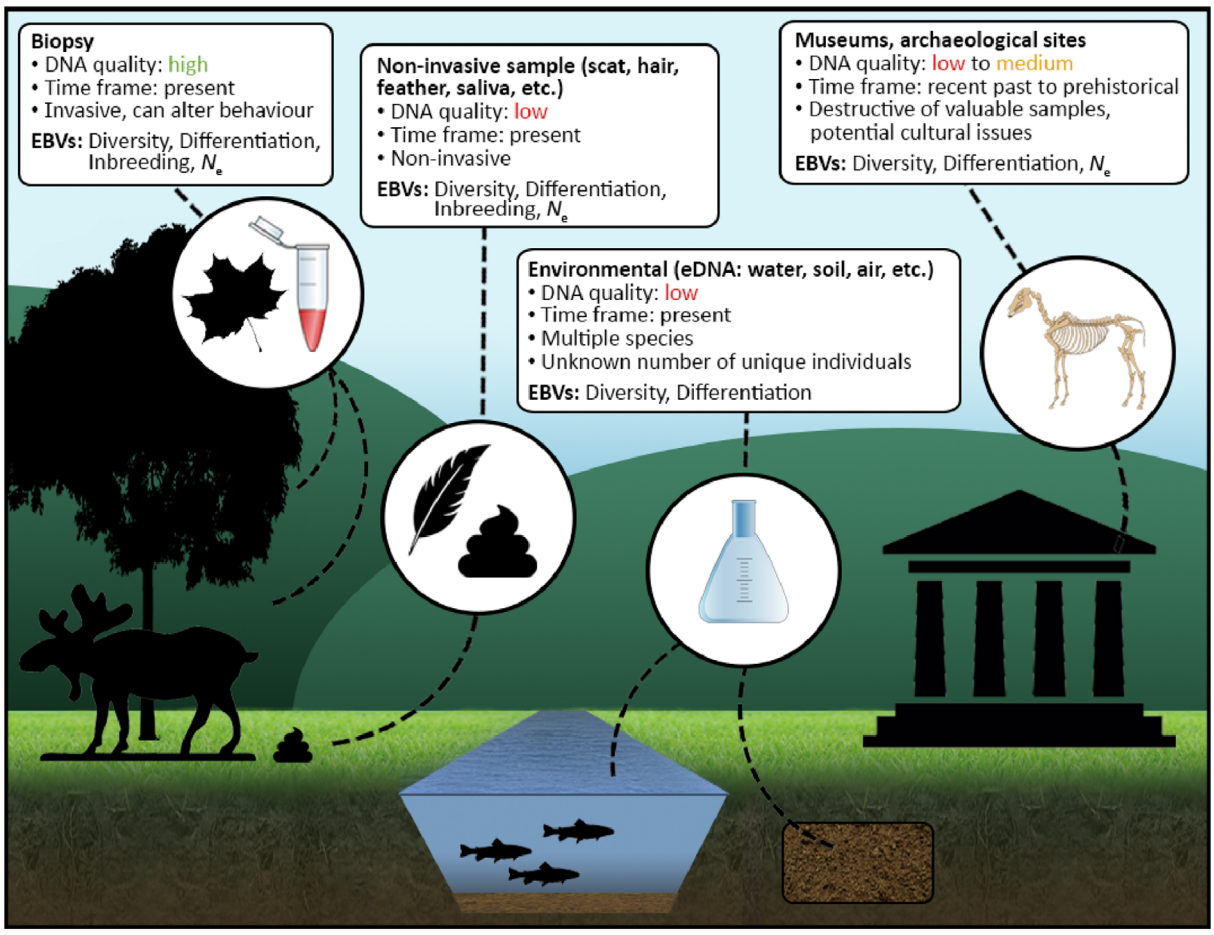


Fig. 5. Sources of DNA for genetic analyses. Genetic material may be obtained directly from tissue samples from extant populations (biopsy or non-invasive), biological collections (e.g. museums), or sub-fossils (sometimes called ancient DNA) or from the environment (i.e. environmental DNA, eDNA). Older samples may have low-quality and low-quantity DNA, restricting the use of certain Genetic Essential Biodiversity Variables (EBVs); eDNA is challenging to use for Genetic EBVs since the DNA is typically of lower quality and quantity. These examples indicate information typically assessed with these types of data, and do not represent all possibilities. N_e, effective population size.

comprehensive analyses. Advancing the necessary supporting components, such as computing infrastructure and bioinformatics, will facilitate data accessibility and application.

As for many biodiversity observations, genetic composition data can be obtained from different types of scientific studies (e.g. assessing the conservation status of important species; testing ecological hypotheses; stock assessments in forestry or fisheries; taxonomy; forensic investigations). The spatial extent of studies is highly variable; sampling often focuses on a specific region but sometimes includes a species' entire range. It is critical that samples are collected in an observational design that considers the expected data product, such as prediction of possible future change under different scenarios; nonetheless numerous studies use opportunistic and biased sampling. If the goal is to assess temporal trends, sampling effort should be comparable between time points. Most studies are from a single time point, but some are longitudinal. For example, many national agencies and research institutions sample fish stocks, large mammals (e.g. lynx, wolves, bears), and forest trees regularly to monitor genetic richness, genetic differentiation, and N_e. One example is the Columbia River Basin in the USA where state, federal, tribal, and non-profit organizations have genetically monitored multiple salmonid species and populations over several decades with yearly funding of hundreds of millions of USD (Hand *et al.*, 2018). This system is one of the most heavily monitored in the world because of socioeconomic and cultural factors coupled with the critical ecological importance of the basin.

Where the study intent is to calculate Genetic EBVs, ideally each step in the data workflow should be guided by standard operating procedures (SOPs) that are feasible, cost-effective and publicly available to produce an 'EBV-ready data set' (Kissling *et al.*, 2018) from which indicators can be derived. SOPs will ensure that EBVs are measured in consistent ways, increase data set interoperability and re-

use, and expand the uptake of Genetic EBVs for diverse users. Examples of SOPs are: site choice and sampling methodology (i.e. choice of individuals, relevant spatial scales and temporal frequency), environmental and biological observations to collect, methods of tissue preservation, data quality control, software/ analysis pipelines, and data archiving and metadata storage to ensure alignment with FAIR principles (Findable, Accessible, Interoperable, and Reproducible). Several communities have developed genetic SOPs. One of the best known is the International Whaling Commission's assessment of stock structure which, in cooperation with associated scientists, has developed best practices for sampling genetic data, data analysis, reporting genetic variation measures including N_e over time, defining population units, and assessing genetic data quality (IWC, 2018; Waples *et al.*, 2018). Other research initiatives have established field, laboratory, sequencing and analysis SOPs appropriate for their taxonomic group, and/or genetic methodology [e.g. Oz Mammals Genomic Initiative (Eldridge *et al.*, 2020); Ira Moana Project (Liggins, Noble & the Ira Moana Network, 2021b)]. The Marine Biodiversity Observation Network (MBON) is working to develop best practices for 'omics' information (Goodwin *et al.*, 2019), including the Global Omics Observatory Network (GLOMICON – https://sites.google.com/view/glomicon/home).

(2) Genetic data archiving practices

It is becoming standard practice to deposit genetic data into searchable and open access repositories – this is mandatory in several leading journals (Joint Data Archiving Policy (https://datadryad.org/docs/JointDataArchivingPolicy.pdf). Successful re-use and integration of these data sets, such as to calculate Genetic EBVs, depends on the archive, its interoperability, and the metadata requirements (Section III.3). The International Nucleotide Sequence Database Collaboration repositories [INSDC; Cochrane *et al.*, 2016; primarily a consortium of US, European and Japanese national databases, including the National Center for Biotechnology Information (NCBI)] are the largest publicly available nucleotide databases. These databases host raw DNA sequence data in standard, interoperable formats and they can be queried and programmatically accessed (e.g. 'rentrez' R package, 'geomedb' R package). While standardized and easily accessible, the genetic data deposited into INSDC does not always include the minimum information required to re-create data sets appropriate for the calculation of Genetic EBVs. Archiving practices may also be inconsistent – the deposited DNA sequences are typically only the new, unique sequences derived from the study rather than all observed sequences, and do not provide frequency information that is needed to calculate EBVs reliably [see Section IV (Pope *et al.*, 2015a; Paz-Vinas *et al.*, 2021)]. For processed genetic data, such as microsatellite and SNP data sets, numerous other open access archives are used, including Dryad, FigShare, and GitHub. Taxonomic and region-specific archives also exist (e.g. CartograTree, DIPNet, see Sections III.3 and IV.1), although it is not always easy to find them. Also, these repositories do not always require standard data file formats or have minimum metadata requirements and therefore these data are not easily searched or programmatically accessed.

Most archived genetic data cannot be operationalized yet. Typically, less than 10% of sequences are georeferenced, and most cannot be geolocated based on text descriptions (Gratton *et al.*, 2017; Miraldo *et al.*, 2016; Toczydlowski *et al.*, 2021). For example, Theodoridis *et al.* (2020) examined mammal mitochondrial genetic diversity globally, comprising >150,000 mitochondrial DNA sequences from BOLD and GenBank. They attempted to assign geographic coordinates to sequences without precise geolocation data using GeoNames.org (http://api.geonames.org), but could do so for just 36% of sequences. In short, even for a very well-studied group (mammals), a majority of data could not be used. In addition, most genetic data sets are not directly linked to non-genetic global biodiversity databases (e.g. the Global Biodiversity Information Facility, GBIF), or environmental databases including widely used satellite and *in situ* data repositories, although links to these can be made through sufficient spatial data (Kissling *et al.*, 2018). In some cases, the exact georeferences of the sampled organism or population may be purposely withheld, to protect the study organism (e.g. an endangered species, or highly desirable species) or at the request of Indigenous peoples for reasons of cultural sensitivity regarding this information. Nonetheless, standardized practices of quality control and assurance when generating data and metadata, and during data archival, would help to avoid issues with compounding errors during data aggregation.

(3) Metadata

Metadata associated with archived samples, DNA, sequences, or other genetic data are essential for enabling re-use of these resources for calculating and interpreting EBVs. Ideally, genomic records should contain environmental information or a link to an environmental database with high-quality information about local conditions of the sampled organism. At the specimen level, metadata may include collection date, taxonomic designation, morphological measurements, latitude/longitude, age, condition/health, spatial scale, and reproductive status. At the site or population level (by convention termed 'occurrence' or 'event' level) these may include habitat description or ecoregion classification, co-occurring species, sampling methodology, measured environmental variables (light, soil moisture, etc.), and uncertainty in the geolocation. Metadata regarding the laboratory, sequencing, quality control, and bioinformatic steps may also be recorded, and are important for generating comparable genetic data but are difficult to document and reproduce in a standard way to derive the genetic data set from the raw data. As with laboratory protocols, open text fields for detailed description may be needed.

Despite disparate study objectives, the research community has been active in developing standardized metadata

fields and vocabularies to capture ecological and environmental data for a sample and/or field sampling event, and laboratory and sequencing methodologies. These metadata standards are supported by the Biodiversity Information Standards community [administering the Darwin Core standard (https://dwc.tdwg.org/; Wieczorek *et al.*, 2012)] and the Genomics Standards Consortium (https://press3.mcs.anl.gov/gensc/mixs/; Wooley, Field & Glöckner, 2009; Yilmaz *et al.*, 2011), and operationalized through the Genomic Observatories Metadatabase (GEOME; https://geome-db.org/; Deck *et al.*, 2017, Riginos *et al.*, 2020). If widely adopted, standard metadata fields alongside deposited genetic data would ensure that they are uniformly queryable according to their machine readable metadata (using a web interface or programmatically using the 'geomedb' R package) and interoperable with other data aggregators (i.e. INSDC, GBIF).To help coordinate the use of a minimum set of mandatory and/or recommended metadata fields, GEOME's infrastructure enables 'Teams' and 'Projects' to design their own metadata templates and rules, based on existing fields and vocabularies (including options to maintain the privacy of some fields, such as exact georeferences for aforementioned reasons). This platform helps users coordinate research projects, ensures the interoperability of data sets, allows discovery of data sets, and will help facilitate the calculation of Genetic EBVs.

Although well developed, uptake and use of the standardized metadata fields and vocabularies by the research community will take time as many researchers remain unaware of suggested metadata fields and large-scale environmental data sets to link to their genetic data. Nonetheless, with the support of leading journals (see Sibbett, Rieseberg & Narum, 2020) and pressure from funding agencies to ensure data is searchable, open access and interoperable, adherence to metadata standards and the use of provided infrastructures will increase, helping ensure these data sets can be used to calculate Genetic EBVs. GEO BON will continue to work with the IUCN, the Society for Conservation Biology, and the EU COST Action G-BiKE to promote broader understanding of these practices in the community.

IV OPERATIONALIZING EBVs: FROM SINGLE STUDIES TO GLOBAL USE

In order to better operationalize EBVs, GEO BON has adopted the data structure of a multi-dimensional array as a basic format to represent EBVs (Jetz *et al.*, 2019). Each cell contains an observation or model-based output with the dimensions as species–space–time, at a minimum. There is an expectation that eventually such data structures, termed 'data cubes', will have regional to global coverage, across critical timescales (yearly, decadal, centuries). As for most biodiversity data, genetic surveys typically sample a limited range of space and timescales, resulting in sparse and spatially biased data structures. This section covers two vital steps towards operational EBVs: how large numbers of genetic data sets can be compiled and analyzed across space and time, and how modeling efforts are helping to fill in the sparse genetic data landscape.

(1) Visualizing EBVs in space and time across species

The well-established population genetic theory describing intrinsic drivers of genetic composition, the relatively small number of genetic statistics, and the community commitment to data archiving have led to numerous compilations of genetic data sets and summary statistics at large spatio-temporal scales. Collectively, this effort has paved the way for large-scale studies to describe regional and global patterns of the four proposed Genetic EBVs and their drivers of change. These large-scale macrogenetic studies (Blanchet *et al.*, 2017; Leigh *et al.*, 2021) can be classified into three categories described below (also see Table 2). Each category represents a major step towards integrating observations into an EBV data object and operationalizing EBVs; they also help develop the knowledge and boundary conditions for simulation models and forecasting of genetic diversity (Section IV.2).

The first category of macrogenetic studies involves integrating data sets from multiple species (tens to thousands) spanning large geographic scales (thousands of square kilometres to global coverage) with extensive environmental databases to identify spatial trends or patterns in EBVs (reviewed in Leigh *et al.*, 2021). A study may test for correlations between genetic diversity and macro-ecological correlates such as latitude (Miraldo *et al.*, 2016), assorted environmental variables (Manel *et al.*, 2020), anthropogenic impact (Millette *et al.*, 2020; Schmidt *et al.*, 2020), species diversity (Theodoridis *et al.*, 2020), or life-history traits (De Kort *et al.*, 2021). Such studies may generate multi-species genetic data from scratch (e.g. Taberlet *et al.*, 2012) and/or extract large amounts of data from the literature or from public repositories like GenBank or BOLD (Leigh *et al.*, 2021). For example, Manel *et al.* (2020) repurposed 50,588 mitochondrial DNA barcode sequences [COI (cytochrome c oxidase subunit I) gene] from 5426 fish species to explore global determinants of freshwater and marine fish species' Genetic Diversity EBV (Table 2). The results from these studies can be effectively mapped in the spatial dimension of the GEO BON EBV 'data cube'. However, numerous statistical considerations apply to macrogenetic studies using data extracted from nucleotide databases, such as inconsistent data depositing practices (see Section III and Paz-Vinas *et al.*, 2021). Further, these studies frequently do not assess true 'population' samples, but rather aggregated sequences according to spatial proximity or a grid (e.g. Theodoridis *et al.*, 2020; Millette *et al.*, 2020; Manel *et al.*, 2020; see Paz-Vinas *et al.*, 2021). It is also often hard to reproduce these studies' results because the data sets are extensively modified once they are obtained from archives (Pope *et al.*, 2015b). Macrogenetic studies based on genetic

Table 2. Examples of three broad categories of large-scale studies that establish a foundation to operationalize Essential Biodiversity Variables (EBVs): taxa examined, data source, EBVs compiled, size of data set, and DNA marker type (mtDNA, mitochondrial DNA; microsat, microsatellite; AFLP, amplified fragment length polymorphism). Category 1 = large scale spatial patterns; Category 2 = temporal change; Category 3 = quantitative relationship between driver and Genetic EBV response

Study	Category	Taxa	Data source	Genetic EBV	Number of species	Marker type
Manel *et al.* (2020)	1	Freshwater and marine fishes	GenBank	Genetic Diversity: richness	5426	mtDNA
Lawrence *et al.* (2019)	1	Terrestrial vertebrates	From literature	Genetic Diversity: richness and evenness; Genetic Differentiation: differentiation between units	897	Microsat
De Kort *et al.* (2021)	1	Animals and plants	From literature	Genetic Diversity: evenness	727	Microsat and AFLP
Leigh *et al.* (2019)	2	Animals	From literature	Genetic Diversity: richness and evenness	91	Predominantly microsat
Jordan *et al.* (2019)	2	Plants	From literature	Genetic Diversity: richness and evenness	48	Various
Pinsky & Palumbi (2014)	3	Fish	From literature	Genetic Diversity: richness and evenness	140	Microsat
Breed *et al.* (2015)	3	Plants	From literature	Inbreeding	40	Predominantly microsat

data extracted from the literature (e.g. Lawrence *et al.*, 2019) may be more reliable for representing population-level genetic data. For instance, De Kort *et al.* (2021) compiled and mapped genetic evenness (heterozygosity) EBV values for 8386 georeferenced populations (with $N \geq 10$ sampled individuals) from 727 animal and plant species to explore the effects of life-history traits and temperature stability on genetic diversity. Further, literature-based macrogenetic studies allow for the assessment of critical information such as the sampling effort needed to obtain a given number of alleles or sequences – information that typically cannot be assessed for data extracted from public repositories (Paz-Vinas *et al.*, 2021; Leigh *et al.*, 2021), although metadata stewardship is continually improving (see Section III.3; Riginos *et al.*, 2020).

A second category of large-scale macrogenetic studies seeks to document temporal change. Here, researchers analyze time-series observations of the same population or region to quantify changes in genetic composition directly. These time series can come from long-term monitoring, or a combination of contemporary samples and archive, museum or sediment (e.g. sub-fossil) specimens. For example, Leigh *et al.* (2019) compiled data from 91 species and found that genetic diversity in animals (primarily fish, birds, mammals and insects) has declined about 6% globally. Jordan *et al.* (2019) found decreased genetic diversity in 52% of restoration plantings (e.g. revegetation) compared to their seed source, an effect most strongly associated with seed obtained from few source sites. These results contrast with analyses by Millette *et al.* (2020) and Lawrence *et al.* (2019), emphasizing that some large-scale studies may have less power to detect change due to a limited time span, repurposing data from nucleotide databases, challenges in aggregation, and other reasons (reviewed in Leigh *et al.*, 2021; Table 2). Unfortunately, true time series of observations of the same population or region with similar sampling are rare, making it difficult to construct a data object with a comprehensive time dimension that spans broad timescales. As with other EBV classes, statistical models will be essential to fill in spatial and temporal gaps (Section IV.2).

A third category seeks to test the effects of specific environmental or human impacts *via* 'space for time substitution' or 'case–control' (Table 2). This can help to quantify the impact of a specific factor or process on genetic composition. For example, Pinsky & Palumbi (2014) analyzed microsatellite genetic data from 140 species to document that, on average, overharvested fish populations have 12% lower allelic richness genetic diversity than non-harvested fish. Aguilar *et al.* (2008) used data from 102 plant species to reveal significantly lower genetic richness and evenness in plant populations after habitat fragmentation, especially after more than 100 years of fragmentation. These and other examples (e.g. DiBattista, 2008), provide information for models (e.g. a minimum distance among populations beyond which significant genetic differences arise; Durrant *et al.*, 2014) that can help populate cells in an EBV data cube, as explained in Section IV.2.

Such efforts at *post-hoc* data aggregation and mobilization are complemented by coordinated research networks where scientists are compiling their population genetic data sets into curated, standardized archives, supplementary to the general nucleotide databases like GenBank. This involves labour-intensive retroactive application of archiving standards and requirements. For example, the Diversity of the Indo-Pacific Network project compiled more than 200 data sets for Indo-Pacific marine organisms (http://diversityindopacific.net/; Crandall *et al.*, 2019). Similar examples include: Ira Moana

(https://sites.massey.ac.nz/iramoana/), FishGen (www.fishgen.net) and CartograTree (https://treegenesdb.org/ct). Other efforts are planning genetic assessment of many species as a part of a single project, and simultaneously co-sampling them. For example, the IntraBioDiv Consortium sampled 27 alpine plant species with the same strategy, in the same locations, with a common analysis method (https://www.wsl.ch/en/projects/intrabiodiv.html); this data set has been used for many studies (e.g. Taberlet *et al.*, 2012; Hanson *et al.*, 2017). The Oz Mammals Genomics Initiative (https://ozmammalsgenomics.com/) is using standardized methods to generate conservation genomic data sets for 14 threatened Australian mammals (Eldridge *et al.*, 2020; von Takach *et al.*, 2021). All of these projects show that mining databases and the literature, and greater coordination of planning for data interoperability and FAIR practices, has great value for understanding and operationalizing Genetic EBVs at large scales. International groups such as GEO BON are working to network among these groups and share data standards, lessons learned, and best practices for operationalizing EBVs (https://geobon.org/ebvs/working-groups/genetic-composition/).

(2) Predicting Genetic EBVs beyond current observations *via* modeling

To be spatially comprehensive, EBVs in all classes must rely on a combination of observational and modeled data. For instance, genetic composition data are usually collected from a subset of individuals or locations, which necessitates accurate, spatially predictive models for EBV values at unsampled locations or time points. One simple form of genetic data modeling is spatial interpolation to identify patterns such as clines and hot spots across numerous taxa (Murphy *et al.*, 2008; Paz-Vinas *et al.*, 2018). At large scales, interpolation has been used to provide a global 'picture' of genetic status (Vandergast *et al.*, 2011; Miraldo *et al.*, 2016).

Another form of genetic data modeling aims at establishing a statistical relationship between Genetic EBV measurements and specific environmental variables including climate, habitat and biotic interactions (Hand *et al.*, 2015; Jaffé *et al.*, 2019; Carvalho *et al.*, 2019, 2020) or EBVs in other classes like species diversity (Taberlet *et al.*, 2012; Manel *et al.*, 2020). The simplest approach is to determine a relationship between genetic differentiation and Euclidean distance, known as isolation by distance or genetic spatial autocorrelation. This can help infer a minimum distance beyond which populations have significant genetic differentiation (Durrant *et al.*, 2014, Silva *et al.*, 2020). More complex genetic–environmental relationships (e.g. landscape genetics) can identify potential drivers and correlates of change in genetic variation. These analyses project genetic variation into unsampled locations for which data on the driver are available, using the driver as a surrogate of genetic variation (Wasserman *et al.*, 2012; Hanson *et al.*, 2017). This has practical utility for highlighting migration corridors to limit genetic erosion (DiLeo & Wagner, 2016; Monteiro *et al.*, 2019), identifying priority conservation areas aimed at protecting within-species genetic composition [e.g. conservation planning (Paz-Vinas *et al.*, 2018; Hanson *et al.*, 2020)], or modeling population vulnerability and changes in adaptive capacity due to environmental change (Fitzpatrick & Keller, 2015; Martins *et al.*, 2018; Razgour *et al.*, 2018; Vranken *et al.*, 2021). For example, Bay *et al.* (2018) identified 'genomic vulnerability' to climate change in populations of the North American yellow warbler (*Setophaga petechia*) based on the mismatch between current genomic variation in climate-adapted loci and predicted future conditions.

Modeling approaches have also been used to combine genetic data with species distribution models. Within-species genetic lineages (the Genetic Differentiation EBV) are increasingly used to improve ecological niche model (ENM) predictions of species distributions. ENMs constructed for within-species genetic units offer a more accurate picture of past and present spatial distribution of a species (e.g. Shinneman *et al.*, 2016), as well as future distributions given adaptive variation and potential climatic changes (e.g. Maguire *et al.*, 2018; Razgour *et al.*, 2019), which helps assess vulnerability of a species and possible changes to genetic composition. Meanwhile, ENMs are also used to assess statistical relationships between environmental suitability and heterozygosity, which could be used to map predicted genetic diversity across continuous space (Diniz-Filho *et al.*, 2015).

One remarkable form of genetic modeling is to infer EBV values in the past. Because an individual genome contains information on its 'ancestral' genomes, numerous well-developed analytical methods can be used to infer N_e at given time points in the past [e.g. skyline plots, approximate Bayesian computation (ABC), Pairwise Sequentially Markovian Coalescent (PSMC) models] (Mather, Traves & Ho, 2020). For example, Nater *et al.* (2015) used modern DNA and ABC to trace the demographic history of orangutan populations, identifying severe N_e declines in response to changes in the extent of rainforest habitats and hunting by early humans. These methods can help establish pre-impact baselines and natural levels of variability in Genetic EBVs, populate the three-dimensional EBV object (e.g. 'data cube'), and better quantify the impact of drivers through time.

Methods are also well developed for forecasting Genetic EBVs, such as changes in richness due to future habitat change (Alsos *et al.*, 2016). Established population genetic theory (Section II.1*a*), allows for demographic–genetic simulations (see Hoban, 2014) to assess genetic composition change under drivers such as management strategies (e.g. promoting gene flow among populations, harvest, minimizing mortality in specific regions). Such models can also evaluate the efficiency of different monitoring programs and rank EBVs for sensitivity (Hoban *et al.*, 2014).

It is true that genetic data are currently scattered spatiotemporally and collected in non-uniform ways that can result in bias, as is true of many ecological data sets and all classes of EBVs [see review on impacts of sampling biases in ecological data sets by Hughes *et al.* (2021); Boakes *et al.*, 2010; Fithian

et al., 2015; Hortal *et al.*, 2015; Jetz *et al.*, 2019; Meyer *et al.*, 2016]. Although genetic data are relatively sparse across landscapes, the aforementioned approaches are filling gaps in genetic information and predicting EBVs spatially and temporally, and across multiple species. Further, the development and uptake of these proposed EBVs could improve global standardized data collection, including use of archive repositories for genetic data and uniform metadata. The existing limitations will continue to be overcome through combining additional supporting measures, such as through statistical, modeling, bioinformatic and software guidelines. To help guide and advance genetic data collection for restoring and conserving biodiversity, guidance is being developed to encourage efficient progress and rapid uptake of genomics for spatial conservation planning (e.g. Kershaw *et al.*, 2021) and natural resource and functional ecosystem restoration (e.g. Mohr *et al.*, 2022). A strong focus on education, training and outreach of practitioners and decision makers, especially in relation to operationalizing EBVs, is also needed (Wasserman *et al.*, 2010; Brown *et al.*, 2016).

V FUTURE DIRECTIONS FOR GENETIC EBVs

(1) What is needed for Genetic EBV operationalization?

A number of challenges exist for Genetic EBV operationalization. These include standardizing bioinformatic pipelines, standardizing sampling for adequate characterization of each EBV, cost, scaling issues, and choosing target species. As Navarro *et al.* (2017, p. 161) explain, the EBV framework as a whole also 'faces challenges emerging from the lack of global monitoring schemes, the integration of datasets resulting from different collection methods, and technical issues related to data product structure, storage, workflow execution, and legal interoperability'. In Genetic EBVs, different spatial and temporal resolution and measurement units need to be harmonized and standardized to allow for integration of data sets from different sources.

For greater adoption of Genetic EBVs by the community, it is critical to demonstrate their value (as we attempt to achieve here), and provide education, training, outreach, and guidance, and also develop necessary supporting components, such as bioinformatics and computing infrastructure. Further, clear protocols and decision trees providing guidance on particular methods or software to use for a given problem, question, or data set (while being mindful of their caveats) could improve uptake. Last, adequate funding to support these goals can allow for increased and sustainable utilization. We and others are continuing to make progress in these areas and are committed to focus on these goals. With these in place, the likelihood of Genetic EBV application should increase.

(2) Challenges: scale, standardization and costs

There is a general lack of systematic global biodiversity data collection and monitoring. Similar to populations and traits, monitoring of genetic change will require combining data sets collected from different species in different locations across time and with different methods – a substantial statistical challenge (see Section IV). Comparisons of Genetic EBVs across taxa will also need to account for correlations with phylogenetic or trait variation (Romiguier *et al.*, 2014).

Just as species or traits may be observed through a variety of measurements, each with its own error rates and biases, genetic information is measured with different markers such as microsatellites, SNPs, or DNA sequences, and there is a need for harmonization across these measurements, particularly to incorporate temporal change. For example, it is not currently possible to compare Genetic EBV values from microsatellites to values from SNPs, without data on both marker types for the same samples. Looking forward, most data sets will soon be sequence-based (e.g. candidate gene resequencing, whole-genome resequencing, reduced-representation sequencing) which are likely to be more robust for long-term monitoring as the data are forward-compatible (see Galla *et al.*, 2019; Wright *et al.*, 2019). Additionally, standardized quality control and assurance practices could ensure that the large volumes of data and metadata harmonized for EBVs do not lead to 'small sources of systematic or random error [that] can cause spurious results or obscure real effects' (Laurie *et al.*, 2010, p. 591).

Another challenge is a lack of standardized minimum sampling size and spatiotemporal sampling guidelines (although a single strategy is impossible as it must adapt depending on organisms and locations), although some recognized sampling rules-of-thumb have been established and validated [e.g. Hale, Burg & Steeves (2012) for microsatellites; Nazareno *et al.* (2017) for SNPs]. A larger challenge is likely a current lack of standard bioinformatic pipelines and decision-making guidelines, e.g. standard thresholds for filtering genetic data. Genetic differentiation and genetic diversity in reduced-representation sequencing data sets can be affected by bioinformatic processing such as SNP calling (e.g. Wright *et al.*, 2019; Graham *et al.*, 2020). SNP calling will depend on whether a reference-based *versus de novo* approach is used, with reference-based approaches being preferred when possible (Shafer *et al.*, 2017; Luikart *et al.*, 2018). The rapidly decreasing cost of whole-genome resequencing will eventually diminish these problems, however. While there will not be a single analytical pipeline that will be appropriate for calculating all Genetic EBVs for all species, having a set of standard methods that are agreed upon, helpful guides for key choices, and standards and templates for documenting informatic pipelines (with well-understood values and rules), will aid with transparency, repeatability, and integration (Tahsin *et al.*, 2016, Gratton *et al.*, 2017, Duruz *et al.*, 2019).

A final challenge is the high cost of gathering and processing genetic data. DNA extraction and sequencing are expensive (*c.* US$10–50 per sample for laboratory supplies

and US$10–100 per sample for personnel time) and require specialized equipment and personnel training. Funding and opportunities for laboratory work and data analysis are limited or absent in some regions, including in many countries within critical biodiversity hotspots. An additional challenge is choosing appropriate representative species and ecosystems to monitor with genetic data. The best choice will be species for which a change in genetic variation serves as a signal of change in other species (indicator species). When efforts and funding are constrained, we suggest that a few representatives from each group and ecosystem be chosen, including foundation/keystone species. Species selection should also consider the amount of pre-existing 'genomic resources', such as annotated reference genomes for the genus. Species can also be selected for genetic monitoring based on the representation of key areas: conservation priority, cultural importance, ecosystem services, wild harvesting, and important game species (as in Hollingsworth *et al.*, 2020). The provisioning of sample access, and uptake and application of EBVs, will be determined by conservation practitioners, including Indigenous peoples as stewards of the natural environment and genetic resources. The Nagoya Protocol (https://www.cbd.int/abs/) to the CBD acknowledges the link of biodiversity conservation with Indigenous peoples. In many nations, Indigenous peoples will be responsible for authorizing access to genetic samples and derived data. Geneticists need to work closely with Indigenous communities (McCartney *et al.*, 2021) to ensure sharing of benefits from research activity and aligning research with Indigenous worldviews, such as by following the CARE Principles [Collective Benefit, Authority to Control, Responsibility and Ethics (https://www.gida-global.org/care; Liggins, Hudson & Anderson, 2021a); Carroll *et al.*, 2020)].

(3) Examples of integrating Genetic EBVs into legislation or policy at large scales

National conservation laws and policies, such as the U.S. Endangered Species Act (ESA), seek to preserve biodiversity through the conservation of individual species (Drozdowski, 1995; Waples *et al.*, 2013). Under the ESA, those species, subspecies, or 'distinct population segments' (DPSs) are listed when they are determined to be 'endangered' (in danger of extinction), or 'threatened' (likely to become endangered in the foreseeable future). Although these policies are often criticized for the lack of quantitative metrics in their language for guidance (Waples *et al.*, 2013), inferences from genetic analyses are frequently utilized to identify units to conserve, determine the level at which they should be listed, and evaluate listed species for delisting (Fallon, 2007; Funk *et al.*, 2019). Although there have been efforts to streamline this process and the listing criteria, genetic data are used inconsistently (Kelly, 2010; Coates, Byrne & Moritz, 2018). The establishment of internationally recognized and easily understood Genetic EBVs will provide quantitative metrics making ESA listing more objective and transparent. Globally, Key Biodiversity Areas now integrate genetic diversity and differentiation in their declaration (IUCN, 2016).

Another national initiative is Scotland's pragmatic assessment of genetic composition in representative species *via* 'scorecards'. The scorecards use structured expert reviews to assess several categories of risks: loss of diversity (through population declines, loss of functional variation and loss of divergent lineages); maladaptive hybridization; and low reproduction or constraints on adaptation. The scorecards also assess *ex situ* conservation representation (zoos, seed banks, botanic gardens) and current conservation actions. For each species, there is a rating of risk (along with an evaluation of uncertainty) and of the effectiveness of mitigation. A similar scorecard approach was used in Beckman *et al.* (2019) which focused on all U.S. species of *Quercus* (oaks), an emblematic, well-known genus of ecological, socioeconomic and cultural importance. Scorecard approaches, including genetic issues and genetic data, could easily be adapted to include Genetic EBVs.

The European Union (EU) Habitats Directive is the centerpiece of biodiversity conservation legislation in the Union. It focuses both on the preservation of habitats and species of conservation concern and requires each member state to define and pursue favorable reference conditions for habitats and species. Since 2012, the directive explicitly mentions that the criteria for a favorable reference population should include genetic information (population structure, genetic connectivity, etc.), and that this favorable reference population is by definition larger than the minimum viable population size (Evans & Arvela, 2011; Laikre *et al.*, 2009). However, general guidelines on how populations should be defined are not provided and are interpreted differently by each member state. Several member states are working to generally implement genetic criteria in their evaluations (Mergeay, 2012), including the N_e EBV. Other EU initiatives that incorporate genetic approaches, and could benefit from EBVs include the UK Strategy for Forest Genetics Resources (Trivedi *et al.*, 2019) and EUFORGEN (de Vries *et al.*, 2015).

(4) Future directions and opportunities

We have focused on the use of genetic data sets, and models that build on such data sets, for EBVs. Nonetheless, large parts of the world lack genetic data observations, necessitating a discussion of proxies or substitutes for genetic data. Suitable proxies could also overcome challenges surrounding the cost of obtaining genetic data. While there are currently no proxies that can fully replace genetic sampling to obtain Genetic EBVs, some show promise, including geographic range loss, decrease in number of adult individuals, or fragmentation rates (see Section IV.2). Decades of theoretical work and hundreds of studies have shown strong genetic composition responses to these proxies as they are connected to basic genetic processes (Sections II.1*a* and IV). For instance, Khoury *et al.* (2019) advocate that safeguarding genetic diversity can be based on the percentage of geographic range found in protected areas (*in situ*) or represented

in seed banks (*ex situ*). However, genetic composition responses are mediated through species- and situation-specific factors, and responses are highly non-linear, such that this relationship may be only weakly predictive (Alsos *et al.*, 2012). Future research effort is needed towards establishing statistical relationships with proxies, identifying how species traits mediate these relationships, and developing statistical frameworks and models (see Section IV.2). Simulations can test how well proxies relate to genetic diversity, i.e. the simulation of loss of range size. Another potential proxy is pedigree data or studbooks, which allow direct inference of N_e, relatedness/inbreeding, and loss of genetic evenness (heterozygosity) *via* available software (e.g. PMx), although this is only applicable to domesticated species, zoo populations, and a few highly monitored natural populations.

The genetic data obtained for Genetic EBVs can also provide information useful to other EBV classes (Gurgel *et al.*, 2020). Genetic data can be used to infer population size (either by mark–recapture or *via* N_e) for the Species Abundance EBV in the Species Populations class, or to delimit species occurrence (Species Population or Community Composition EBVs) through eDNA sampling (see Section I.3). Coordination of sampling is another useful synergy; genetic sample collection could occur concurrently with Species Populations and Traits EBV data collection to facilitate cost-effectiveness and allow cohesive assessments across multiple EBVs, an important consideration for Biodiversity Observation Networks (BONs). There are opportunities across EBV classes to coordinate development of standardized sampling protocols, data-storage structures, and modeling and harmonization techniques. Lastly, we have little knowledge of how Genetic EBVs and other EBV classes covary in response to environmental change (although see Razgour *et al.*, 2019), including whether EBV classes exhibit related changes to response or response time (e.g. lags in response). Indicators or other high-level reporting metrics could also integrate multiple classes to illustrate a comprehensive picture of biodiversity.

The final phase of EBV operationalization is policy implementation. EBVs have great potential to form the basis of indicators for policy (Navarro *et al.*, 2017). Unfortunately, currently recommended CBD indicators do not track within-species genetic variation and are not built upon Genetic EBVs (Appendix S4; Laikre *et al.*, 2020; Hoban *et al.*, 2020). Still, there are emerging examples of Genetic EBVs or similar concepts in use in policy frameworks that may inform and engender further efforts across the globe (see Section V.3).

Genetic and genomic methods continue to advance rapidly, creating opportunities for refining Genetic EBVs in the future, particularly focused on quantifying uncertainties and creating open data frameworks. Decreasing costs and ongoing research efforts are rapidly increasing the amount and quality of reference genome data available, and knowledge of genomic processes, genes and their functions. Long-read sequencing technologies like PacBio and Oxford Nanopore technology will soon allow monitoring *via* whole-genome sequencing for species of small to moderate genome size. To facilitate such studies, tissue samples should be stored in biorepositories, e.g. biobanks, for use in future advanced analyses. Developments in transcriptomics and metabolomics will lead to advances in mechanistic models and prediction of specific genomic responses to environmental change. Technological advances such as synthetic biology [e.g. CRISPR (clustered regularly interspaced short palindromic repeats) modification] may lead to transformative ways to manage and conserve genetic variation. Additionally, rapid sequencing technology for use 'in the field' promises to allow near-immediate genetic data collection for urgent questions (e.g. for prioritizing harvest from specific source populations in fish and in wildlife management). Rapid, 'pocket-size' genomic sequencers such as the Nanopore cost less than US$1000 and plug into a laptop. Although genetic composition is not currently amenable to remote-sensing technologies such as satellites, autonomous collection *via in situ* recording infrastructure (such as the U.S. National Ecological Observatory Network, NEON), hair traps, or automated sampling devices may soon allow DNA capture, sequencing and analysis in the field (Turon *et al.*, 2020). Drones with increased autonomy allow sampling in elusive species or inaccessible sites (e.g. cliffs, oceans), and regularly repeated sampling. Advances in eDNA analysis are also anticipated, although due to eDNA quantities and degradation, it might not be possible to use eDNA for all EBVs (Section I.3).

Technology for collecting environmental data that are necessary for interpreting and modeling genetic data are also advancing. 'Biologging', a technique for live registration of physiological parameters, coupled with epigenetics, will help to determine the importance of genetic variants for individual phenotypes and ecosystem functions (e.g. Tree Talker© is a micro-device for measuring water transport in trees, diametrical growth, and spectral characteristics; Valentini *et al.*, 2019). Coupling these data with Genetic EBVs will help to determine underlying causal factors behind genetic, species, and ecosystem diversity, enable us to predict future responses, and help to identify adaptive genetic variants that can respond to climate and environmental change.

VI CONCLUSIONS

Here we summarize our review and describe concrete steps for advancing operationalization of Genetic EBVs.

(1) We propose four Genetic EBVs: (*i*) Genetic Diversity, (*ii*) Genetic Differentiation, (*iii*) Inbreeding, and (*iv*) Effective Population Size. These EBVs offer a viable means of monitoring the essential characteristics of genetic composition, genetic health, and genetic viability of wild and domesticated populations and species, at regional to global scales.

(2) The technical and analytical foundation of these Genetic EBVs is well developed, and workflows for

archiving, obtaining and analyzing genetic data are maturing. There are highly standardized data structures and file formats for storage, and well-developed open-access pipelines for calculating standard statistics underlying the EBVs. Numerous statistical and computational developments have led to highly advanced models for inferring and predicting genetic change, which are complemented by increasingly large-scale genetic data set aggregation.

(3) Despite the robust underpinnings of these Genetic EBVs, additional developments and research efforts are required. Specifically, progress towards increased standardization and shared resources will lead to routine, robust uptake and operationalization of Genetic EBVs in biodiversity assessments and policy. The next steps are:

(*a*) Improved and upscaled monitoring of temporal changes in genetic composition. Standardized temporal measures of genetic composition remain scarce due to limited longitudinal genetic monitoring, biases in location and taxonomy of the available data, and still unsolved statistical challenges in aggregating and harmonizing existing genetic data sets in large-scale analyses.

(*b*) Standardization of genetic methods. Standard laboratory and informatic methodologies, and guidelines for consistent, routine archiving and reuse of genetic data, still require improvements to make all genetic data interoperable.

(*c*) Better recording and use of genetic metadata. There are broad opportunities for increased use of genetic archival databases and biological and environmental metadatabases, and for better links between genetic and environmental repositories and observatories to improve data searching and inferences regarding the drivers of changes in Genetic EBVs. There is a need to develop better proxies in the absence of genetic data.

(*d*) Development of genetic resources and technology. Further work towards development of key genetic resources (reference genomes), especially for IUCN threatened species and indicator species and those that underpin biodiversity (foundation species), as well as increased deployment of genetic technologies to additional countries and ecosystems, will greatly aid Genetic EBV calculations.

(4) Genetic EBVs capture the genetic status and trajectory of life on Earth. Progress is being made towards integration of genetic variation observations into major international regulations and legislation. International initiatives are working towards a framework for monitoring Genetic EBVs globally, including GEO BON, the IUCN Conservation Genetics Specialist Group and the EU COST Action G-BiKE.

(5) The EBVs proposed here are centered on genetic population processes and will remain relevant, even as innovative calculation methods are developed and with advances in genomic technologies. Implementation of our recommendations for data curation, sharing and integration, alongside increasingly detailed assessments of Genetic EBVs, can lead to robust assessment of genetic status and improved conservation and management of the world's biomes into the future.

VII ACKNOWLEDGEMENTS

We would like to thank the organizations and groups that have contributed to this project, including NIMBioS, GEO BON, G-BiKE, and IUCN. We would like to acknowledge John Robinson and Sarah Pearson for their early input to this review. This article is based on work from COST Action G-BiKE, CA 18134, supported by COST (European Cooperation in Science and Technology), www.cost.eu. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the United States Government.

VIII. REFERENCES

References identified with an asterisk (*) are used only in the supporting information.

Ackerman, M. W., Hand, B. K., Waples, R. K., Luikart, G., Waples, R. S., Steele, C. A., Garner, B. A., McCane, J. & Campbell, M. R. (2017). Effective number of breeders from sibship reconstruction: empirical evaluations using hatchery steelhead. *Evolutionary Applications* **10**, 146–160.

*Adamack, A. T., Gruber, B. & Dray, S. (2014). PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* **5**, 384–387.

Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y. & Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology* **17**, 5177–5188.

Ahrens, C. W., Rymer, P. D., Stow, A., Bragg, J., Dillon, S., Umbers, K. D. L. & Dudaniec, R. Y. (2018). The search for loci under selection: trends, biases and progress. *Molecular Ecology* **27**, 1342–1356.

*Alexander, D. H., Novembre, J. & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Resources* **19**, 1655–1664.

Almeida-Rocha, J. M., Soares, L. A. S. S., Andrade, E. R., Gaiotto, F. A. & Cazetta, E. (2020). The impact of anthropogenic disturbances on the genetic diversity of terrestrial species: a global meta-analysis. *Molecular Ecology* **29**, 4812–4822.

Alsos, I. G., Ehrich, D., Thuiller, W., Eidesen, P. B., Tribsch, A., Schönswetter, P., Lagaye, C., Taberlet, P. & Brochmann, C. (2012). Genetic consequences of climate change for northern plants. *Proceedings of the Royal Society B: Biological Sciences* **279**, 2042–2051.

Alsos, I. G., Sjögren, P., Edwards, M. E., Landvik, J. Y., Gielly, L., Forwick, M., Coissac, E., Brown, A. G., Jakobsen, L. V., Føreid, M. K. & Pedersen, M. W. (2016). Sedimentary ancient DNA from Lake Skartjørna, Svalbard: assessing the resilience of arctic flora to Holocene climate change. *The Holocene* **26**, 627–642.

*Anderson, E. C. & Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217–1229.

Balázs, G., Vörös, J., Lewarne, B. & Herczeg, G. (2020). A new non-invasive in situ underwater DNA sampling method for estimating genetic diversity. *Evolutionary Ecology* **34**, 633–644.

Balkenhol, N., Cushman, S. A., Waits, L. P. & Storfer, A. (2016). *Landscape Genetics: Concepts, Methods, Applications*. Wiley-Blackwell, Hoboken.

Bay, R. A., Harrigan, R. J., Underwood, V. L., Gibbs, H. L., Smith, T. B. & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**, 83–86.

Beckman, E., Meyer, A., Denvir, A., Gill, D., Man, G., Pivorunas, D., Shaw, K. & Westwood, M. (2019). *Conservation Gap Analysis of Native U.S. Oaks*. The Morton Arboretum, Lisle.

*Beerli, P., Mashayekhi, S., Sadeghi, M., Khodaei, M. & Shaw, K. (2019). Population genetic inference with MIGRATE. *Current Protocols in Bioinformatics* **68**, e87.

Bi, K., Linderoth, T., Singhal, S., Vanderpool, D., Patton, J. L., Nielsen, R., Moritz, C. & Good, J. M. (2019). Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLoS Genetics* **15**, e1008119.

*Biscarini, F., Cozzi, P., Gaspa, G. & Marras, G. (2018). detectRUNS: detect runs of homozygosity and runs of heterozygosity in diploid genomes. R package version 0.9.5. Electronic file available at https://CRAN.R-project.org/package=detectRUNS. Accessed March 10, 2022.

Blanchet, S., Prunier, J. G. & De Kort, H. (2017). Time to go bigger: emerging patterns in macrogenetics. *Trends in Genetics* **33**, 579–580.

Boakes, E. H., McGowan, J. K., Fuller, R. A., Chang-qing, D., Clark, N. E., O'Connor, K. & Mace, G. M. (2010). Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology* **8**, e1000385.

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A. & Zemp, M. (2014). The concept of essential climate variables in support of climate research, applications, and policy. *Bulletin of the American Meteorological Society* **95**, 1431–1443.

Breed, M. F., Ottewell, K. M., Gardner, M. G., Marklund, M. H. K., Dormontt, E. E. & Lowe, A. J. (2015). Mating patterns and pollinator mobility are critical traits in forest fragmentation genetics. *Heredity* **115**, 108–114.

Brown, J. L., Weber, J. J., Alvarado-Serrano, D. F., Hickerson, M. J., Franks, S. J. & Carnaval, A. C. (2016). Predicting the genetic consequences of future climate change: the power of coupling spatial demography, the coalescent, and historical landscape changes. *American Journal of Botany* **103**, 153–163.

Bruford, M. W., Davies, N., Dulloo, M. E., Faith, D. P. & Walters, M. (2017). Monitoring changes in genetic diversity. In *The GEO Handbook on Biodiversity Observation Networks*, pp. 107–128. Springer, Cham.

Brummitt, N., Regan, E., Weatherdon, L. V., Martin, C., Geijzendorffer, I. R., Rocchini, D., Gavish, Y., Haase, P., Marsh, C. J. & Schmeller, D. S. (2017). Taking stock of nature: essential biodiversity variables explained. *Biological Conservation* **213**, 252–255.

Carroll, E. L., Bruford, M. W., DeWoody, J. A., Leroy, G., Strand, A., Waits, L. & Wang, J. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evolutionary Applications* **11**, 1094–1119.

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J. & Hudson, M. (2020). The CARE principles for indigenous data governance. *Data Science Journal* **19**, 1–12.

Carvalho, C. S., Forester, B. R., Mitre, S. K., Alves, R., Imperatriz-Fonseca, V. L., Ramos, S. J., Resende-Moreira, L. C., Siqueira, J. O., Trevelin, L. C., Caldeira, C. F., Gastauer, M. & Jaffé, R. (2020). Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration. *Molecular Ecology Resources* **21**, 44–58.

Carvalho, C. S., Lanes, E. C. M., Silva, A. R., Caldeira, C. F., Carvalho-Filho, N., Gastauer, M., Imperatriz-Fonseca, V. L., Nascimento Júnior, W., Oliveira, G., Siqueira, J. O., Viana, P. L. & Jaffé, R. (2019). Habitat loss does not always entail negative genetic consequences. *Frontiers in Genetics* **10**, 1011.

*CBD (2010). X/2 The Strategic Plan for Biodiversity 2011–2020 and the Aichi Biodiversity Targets. In Decisions Adopted by the Conference of the Parties to the Convention on Biological Diversity at Its Eleventh Meeting, Nagoya, Japan, 18–29 October 2010. Secretariat of the Convention on Biological Diversity. Electronic file available at http://www.cbd.int/doc/decisions/cop-10/full/cop-10-dec-en.pdf. Accessed March 10, 2022.

*CBD (2016). XII/28 Indicators for the Strategic Plan for Biodiversity 2011–2020 and the Aichi Biodiversity Targets. In Decisions Adopted by the Conference of the Parties to the Convention on Biological Diversity at its Thirteenth Meeting, Cancun, Mexico, 4–17 December 2016. Secretariat of the Convention on Biological Diversity. Electronic file available at https://www.cbd.int/doc/decisions/cop-13/cop-13-dec-28-en.pdf. Accessed March 10, 2022.

*Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M. & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7.

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics* **10**, 195–205.

Coates, D. J., Byrne, M. & Moritz, C. (2018). Genetic diversity and conservation units: dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution* **6**, 165.

Cochrane, G., Karsch-Mizrachi, I., Takagi, T. & INSDC (2016). The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research* **44**, D48–D50.

Crandall, E. D., Riginos, C., Bird, C. E., Liggins, L., Treml, E., Beger, M., Barber, P. H., Connolly, S. R., Cowman, P. F., DiBattista, J. D., Eble, J. A., Magnuson, S. F., Horne, J. B., Kochzius, M., Lessios, H. A., et al. (2019). The molecular biogeography of the Indo-Pacific: testing hypotheses with multispecies genetic patterns. *Global Ecology and Biogeography* **28**, 943–960.

Daly, A. J., Baetens, J. M. & De Baets, B. (2018). Ecological diversity: measuring the unmeasurable. *Mathematics* **6**, 119.

*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**, 2156–2215.

De Kort, H., Prunier, J. G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V. M. & Blanchet, S. (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. *Nature Communications* **12**, 516.

de Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology* **56**, 879–886.

de Vries, S. M., Alan, M., Bozzano, M., Burianek, V., Collin, E., Cottrell, J., Ivankovic, M., Kelleher, C. T., Koskela, J., Rotach, P., Vietto, L. & Yrjänä, L. (2015). *Pan-European Strategy for Genetic Conservation of Forest Trees and Establishment of a Core Network of Dynamic Conservation Units*. European Forest Genetic Resources Programme (EUFORGEN), Biodiversity International, Rome. xii, p. 3.

Deck, J., Gaither, M. R., Ewing, R., Bird, C. E., Davies, N., Meyer, C., Riginos, C., Toonen, R. J. & Crandall, E. D. (2017). The Genomic Observatories Metadatabase (GeOMe): a new repository for field and sampling event metadata associated with genetic samples. *PLoS Biology* **15**, e2002925.

DiBattista, J. D. (2008). Patterns of genetic variation in anthropogenically impacted populations. *Conservation Genetics* **9**, 141–156.

Díez-del-Molino, D., Sánchez-Barreiro, F., Barnes, I., Gilbert, M. T. P. & Dalén, L. (2018). Quantifying temporal genomic erosion in endangered species. *Trends in Ecology & Evolution* **33**, 176–185.

DiLeo, M. F. & Wagner, H. H. (2016). A landscape ecologist's agenda for landscape genetics. *Current Landscape Ecology Reports* **1**, 115–126.

Diniz-Filho, J. A. F., Rodrigues, H., Telles, M. P. D. C., Oliveira, G. D., Terribile, L. C., Soares, T. N. & Nabout, J. C. (2015). Correlation between genetic diversity and environmental suitability: taking uncertainty from ecological niche models into account. *Molecular Ecology Resources* **15**, 1059–1066.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J. & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources* **14**, 209–214.

Drozdowski, J. (1995). Saving an endangered act: the case for a biodiversity approach to ESA conservation efforts. *Case Western Reserve Law Review* **45**, 553.

*Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**, 1969–1973.

Durrant, H. M. S., Burridge, C. P., Kelaher, B. P., Barrett, N. S., Edgar, G. J. & Coleman, M. A. (2014). Implications of macroalgal isolation by distance for networks of marine protected areas. *Conservation Biology* **28**, 438–445.

Duruz, S., Sevane, N., Selmoni, O., Vajana, E., Leempoel, K., Stucki, S., Orozco-terWengel, P., Rochat, E., Dunner, S., The NEXTGEN Consortium, The CLIMGEN Consortium, Bruford, M. W. & Joost, S. (2019). Rapid identification and interpretation of gene-environment using the R.SamBada landscape genomics pipeline. *Molecular Ecology Resources* **19**, 1355–1365.

Eldridge, M. D. B., Deakin, J. E., MacDonald, A. J., Byrne, M., Fitzgerald, A., Johnson, R. N., Moritz, C., Palmer, S. & Young, A. (2020). The Oz Mammals Genomics (OMG) initiative: developing genomic resources for mammal conservation at a continental scale. *Australian Zoologist* **40**, 505–509.

Ellegren, H. & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics* **17**(7), 422–433.

Evans D. & Arvela M. (2011). Assessment and reporting under Article 17 of the Habitats Directive. Explanatory Notes & Guidelines for the period 2007–2012. *European Topic Centre on Biological Diversity*. Electronic file available at https://circabc.europa.eu/sd/a/2c12cea2-f827-4bdb-bb56-3731c9fd8b40/Art17-Guidelines-final.pdf.

*Excoffier, L. & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564–567.

*Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V. C. & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics* **9**, e100390.

Fallon, S. M. (2007). Genetic data and the listing of species under the U.S. Endangered Species Act. *Conservation Biology* **21**, 1186–1195.

Fenderson, L. E., Kovach, A. I. & Llamas, B. (2020). Spatiotemporal landscape genetics: investigating ecology and evolution through space and time. *Molecular Ecology* **29**, 218–246.

Ferreira, C. M., Sabino-Marques, H., Barbosa, S., Costa, P., Encarnação, C., Alpizar-Jara, R., Pita, R., Beja, P., Mira, A., Searle, J. B. & Paupério, J. (2018). Genetic non-invasive sampling (gNIS) as a cost-effective tool for monitoring elusive small mammals. *European Journal of Wildlife Research* **64**, 1–15.

Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., Holderegger, R. & Widmer, A. (2017). Estimating genomic diversity and population differentiation–an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics* **18**, 69.

Fithian, W., Elith, J., Hastie, T. & Keith, D. A. (2015). Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in Ecology and Evolution* **6**, 424–438.

Fitzpatrick, M. C. & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* **18**, 1–16.

*Foll, M. & Gaggiotti, O. E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–993.

Frankham, R. (2005). Genetics and extinction. *Biological Conservation* **126**, 131–140.

Frankham, R., Bradshaw, C. J. & Brook, B. W. (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation* **170**, 56–63.

Fraser, D. J. & Bernatchez, L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology* **10**, 2741–2752.

*Frichot, E. & François, O. (2015). LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution* **6**, 925–929.

Funk, W. C., Forester, B. R., Converse, S. J., Darst, C. & Morey, S. (2019). Improving conservation policy with genomics: a guide to integrating adaptive potential into U.S. Endangered Species Act decisions for conservation practitioners and geneticists. *Conservation Genetics* **20**, 115–134.

Funk, W. C., McKay, J. K., Hohenlohe, P. A. & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution* **27**, 489–496.

Gaggiotti, O. E., Chao, A., Peres-Neto, P., Chiu, C. H., Edwards, C., Fortin, M. J., Jost, L., Richards, C. M. & Selkoe, K. A. (2018). Diversity from genes to ecosystems: a unifying framework to study variation across biological metrics and scales. *Evolutionary Applications* **11**, 1176–1193.

Galla, S. J., Forsdick, N. J., Brown, L., Hoeppner, M. P., Knapp, M., Maloney, R. F., Moraga, R., Santure, A. W. & Steeves, T. E. (2019). Reference genomes from distantly related species can be used for discovery of single nucleotide polymorphisms to inform conservation management. *Genes* **10**, 9.

Garner, B. A., Hand, B. K., Amish, S. J., Bernatchez, L., Foster, J. T., Miller, K. M., Morin, P. A., Narum, S. R., O'Brien, S. J., Roffler, G., Templin, W. D., Sunnucks, P., Strait, J., Warheit, K. I., Seamons, T. R., et al. (2016). Genomics in conservation: case studies and bridging the gap between data and application. *Trends in Ecology & Evolution* **31**, 81–83.

Garrido-Cardenas, J. A., Mesa-Valle, C. & Manzano-Agugliaro, F. (2018). Trends in plant research using molecular markers. *Planta* **247**, 543–557.

*Gompert, Z. & Buerkle, C. A. (2010). introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources* **10**, 378–338.

Goodwin, K. D., Muller-Karger, F. E., Djurhuus, A., Zeigler Allen, L., Allen, A. E., McCrow, J. P. & Canonico Hyde, G. (2019). Molecular approaches for an operational Marine Biodiversity Observation Network. In *World Seas: An Environmental Evaluation, Vol. III: Ecological Issues and Environmental Impacts*, Second Edition (ed. C. Sheppard), pp. 613–631. Elsevier.

*Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Electronic file available from http://www.unil.ch/izea/softwares/fstat.html.

*Goudet, J. (2005). hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* **5**, 184–186.

Graham, C. F., Boreham, D. R., Manzon, R. G., Stott, W., Wilson, J. Y. & Somers, C. M. (2020). How "simple" methodological decisions affect interpretation of population structure based on reduced representation library DNA sequencing: a case study using the lake whitefish. *PLoS One* **15**, e0226608.

Gratton, P., Marta, S., Bocksberger, G., Winter, M., Trucchi, E. & Kühl, H. (2017). A world of sequences: can we use georeferenced nucleotide databases for a robust automated phylogeography? *Journal of Biogeography* **44**, 475–486.

Graudal, L., Loo, J., Fady, B., Vendramin, G., Aravanopoulos, F. A., Baldinelli, G., Bennadji, Z., Ramamonjisoa, L., Changtragoon, S. & Kjær, E. D. (2020). Indicators of the genetic diversity of trees – state, pressure, benefit and response. In Michele, B., Riina, J., Evert, T., David, B., Leonardo, G., Stephen, C., Sándor, B., Paul, S. & Judy, L., eds. *State of the World's Forest Genetic Resources – Thematic Study*. FAO, Rome.

*Gunther, T. & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics* **195**, 205–220.

Gurgel, C. F. D., Camacho, O., Minne, A. J. P., Wernberg, T. & Coleman, M. A. (2020). Marine heatwave drives cryptic loss of genetic diversity in underwater forests. *Current Biology* **30**, 1199–1206.e1192.

Hale, M. L., Burg, T. M. & Steeves, T. E. (2012). Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *PLoS One* **7**, e45170.

Hand, B. K., Flint, C. G., Frissell, C. A., Muhlfeld, C. C., Devlin, S. P., Kennedy, B. P., Crabtree, R. L., McKee, W. A., Luikart, G. & Stanford, J. A. (2018). A social–ecological perspective for riverscape management in the Columbia River Basin. *Frontiers in Ecology and the Environment* **16**, S23–S33.

Hand, B. K., Lowe, W. H., Kovach, R. P., Muhlfeld, C. C. & Luikart, G. (2015). Landscape community genomics: understanding eco-evolutionary processes in complex environments. *Trends in Ecology & Evolution* **30**, 161–168.

Hanson, J. O., Marques, A., Veríssimo, A., Camacho-Sánchez, M., Velo-Antón, G., Martínez-Solano, I. & Carvalho, S. B. (2020). Conservation planning for adaptive and neutral evolutionary processes. *Journal of Applied Ecology* **57**, 2159–2169.

Hanson, J. O., Rhodes, J. R., Riginos, C. & Fuller, R. A. (2017). Environmental and geographic variables are effective surrogates for genetic variation in conservation planning. *Proceedings of the National Academy of Sciences* **114**, 201711009.

Henly, L. & Wentworth J. (2021). POST (Parliamentary Office of Science and Technology) brief 41, biodiversity indicators. UK Parliament.

Henry, P., Miquelle, D., Sugimoto, T., McCullough, D. R., Caccone, A. & Russello, M. A. (2009). In situ population structure and ex situ representation of the endangered Amur tiger. *Molecular Ecology* **18**, 3173–3184.

*Hey, J. (2010). Isolation with migration models for more than two populations. *Molecular Biology and Evolution* **27**, 905–920.

Hoban, S. (2014). An overview of the utility of population simulation software in molecular ecology. *Molecular Ecology* **23**, 2383–2401.

Hoban, S., Arntzen, J. A., Bruford, M. W., Godoy, J. A., Rus Hoelzel, A., Segelbacher, G., Vilà, C. & Bertorelle, G. (2014). Comparative evaluation of potential indicators and temporal sampling protocols for monitoring genetic erosion. *Evolutionary Applications* **7**, 984–998.

Hoban, S., Bruford, M., D'Urban Jackson, J., Lopes-Fernandes, M., Heuertz, M., Hohenlohe, P. A., Paz-Vinas, I., Sjögren-Gulve, P., Segelbacher, G., Vernesi, C., Aitken, S., Bertola, L. D., Bloomer, P., Breed, M., Rodríguez-Correa, H., et al. (2020). Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. *Biological Conservation* **248**, 108654.

Hoban, S., Campbell, C. D., da Silva, J. M., Ekblom, R., Funk, W. C., Garner, B. A., Godoy, J. A., Kershaw, F., MacDonald, A. J., Mergeay, J., Minter, M., O'Brien, D., Paz-Vinas, I., Pearson, S. K., Perez-Espona, S., et al. (2021). Genetic diversity is considered important but interpreted narrowly in country reports to the Convention on Biological Diversity: current actions and indicators are insufficient. *Biological Conservation* **261**, 109233.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A. & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist* **188**, 379–397.

Hohenlohe, P. A., Funk, W. C. & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology* **30**, 62–82.

Hollingsworth, P. M., O'Brien, D., Ennos, R. A., Yahr, R., Neaves, L., Ahrends, A., Ballingall, K. T., Brooker, R. W., Burke, T., Cavers, S., Dawson, I. K., Elston, D. A., Kerr, J., Marshall, D. F., et al. (2020). *Scotland's Biodiversity Progress to 2020 Aichi Targets: Conserving Genetic Diversity – Development of a National Approach for Addressing Aichi Biodiversity Target 13 that Includes Wild Species*. Scottish Natural Heritage, Inverness.

Hortal, J., Bello, F., Diniz-Filho, J. A., Lewinsohn, T. M., Lobo, J. M. & Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **46**, 523–549.

Hu, X.-S., He, F. & Hubbell, S. P. (2006). Neutral theory in macroecology and population genetics. *Oikos* **113**, 548–556.

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C. & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography* **44**, 1259–1269.

*IPBES (2019). In *Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (eds E. S. Brondizio, J. Settele, S. Díaz and H. T. Ngo). IPBES secretariat, Bonn. 1148 pages. Supplementary Online Materials, Chapter 3. Electronic file available at http://ipbes.net/sites/default/files/2021-1/GA_chapter_3_supplementary_materials.pdf.

IUCN (2016). *A Global Standard for the Identification of Key Biodiversity Areas, Version 1.0*, First Edition . IUCN, Gland.

IWC (2018). Scientific Committee Handbook: Working Methods of the IWC's Scientific Committee. Commission Meeting Document IWC/67/FA/20. 118pp.

Jaffé, R., Veiga, J. C., Pope, N. S., Lanes, E. C. M., Carvalho, C. S., Alves, R., Andrade, S. C. S., Arias, M. C., Bonatti, V., Carvalho, A. T., de Castro, M. S., Contrera, F. A. L., Francoy, T. M., Freitas, B. M., Giannini, T. C., et al. (2019). Landscape genomics to the rescue of a tropical bee threatened by habitat loss and climate change. *Evolutionary Applications* **12**, 1164–1177.

Jamieson, I. G. & Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution* **27**, 578–584.

Jetz, W., McGeoch, M. A., Guralnick, R., Ferrier, S., Beck, J., Costello, M. J., Fernandez, M., Geller, G. N., Keil, P., Merow, C., Meyer, C., Muller-Karger, F. E., Pereira, H. M., Regan, E. C., Schmeller, D. S., et al. (2019). Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology & Evolution* **3**, 539–551.

*Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405.

Jombart, T., Devillard, S. & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**, 94.

*Jones, O. R. & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**, 551–555.

Jordan, R., Breed, M. F., Prober, S. M., Miller, A. D. & Hoffmann, A. A. (2019). How well do revegetation plantings capture genetic diversity? *Biology Letters* **15**, 20190460.

Jost, L. (2006). Entropy and diversity. *Oikos* **113**, 363–375.

*Jost, L. (2010). The relation between evenness and diversity. *Diversity* **2**, 207–202.

Jost, L., Archer, F., Flanagan, S., Gaggiotti, O., Hoban, S. & Latch, E. (2018). Differentiation measures for conservation genetics. *Evolutionary Applications* **11**, 1139–1148.

*Kalinowski, S. T., Wagner, A. P. & Taper, M. L. (2006). ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes* **6**, 576–557.

Kardos, M., Åkesson, M., Fountain, T., Flagstad, Ø., Liberg, O., Olason, P., Sand, H., Wabakken, P., Wikenros, C. & Ellegren, H. (2018). Genomic consequences of intensive inbreeding in an isolated wolf population. *Nature Ecology & Evolution* **2**, 124–131.

Kardos, M., Luikart, G. & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity* **115**, 63–72.

*Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., Prodöhl, P. A. & O'Hara, R. B. (2013). diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution* **4**, 782–788.

Kelly, R. (2010). The use of population genetics in endangered species act listing decisions. *Ecology Law Quarterly* **37**, 1107–1158.

Kershaw, F., McClintock, W., Andrews, K. R., Riet-Sapriza, F. G., Caballero, S., Tetley, M. J., Notarbartolo di Sciara, G., Hoyt, E., Goldberg, G., Chou, E., Kane-Ritsch, K. & Rosenbaum, H. C. (2021). Geospatial genetics: integrating genetics into marine protection and spatial planning. *Aquatic Conservation: Marine and Freshwater Ecosystems* **31**, 2440–2458.

Kettenring, K. M., Mercer, K. L., Adams, C. R. & Hines, J. (2014). Application of genetic diversity-ecosystem function research to ecological restoration. *Journal of Applied Ecology* **51**, 339–348.

Keyghobadi, N., Roland, J., Matter, S. F. & Strobeck, C. (2005). Among-and within-patch components of genetic diversity respond at different rates to habitat fragmentation: an empirical demonstration. *Proceedings of the Royal Society B: Biological Sciences* **272**, 553–560.

*Khoury, C. K., Amariles, D., Soto, J. S., Diaz, M. V., Sotelo, S., Sosa, C. C., Ramírez-Villegas, J., Achicanoy, H. A., Velásquez-Tibatá, J., Guarino, L., León, B., Navarro-Racines, C., Castañeda-Alvarez, N., Dempewolf, H., et al. (2019). Comprehensiveness of conservation of useful wild plants: an operational indicator for biodiversity and sustainable development targets. *Ecological Indicators* **98**, 420–429.

Kimura, M. & Crow, J. F. (1963). The measurement of effective population number. *Evolution* **17**, 279–288.

Kissling, W. D., Walls, R., Bowser, A., Jones, M. O., Kattge, J., Agosti, D., Amengual, J., Basset, A., van Bodegom, P. M., Cornelissen, J. H. C., Denny, E. G., Deudero, S., Egloff, W., Elmendorf, S. C., Alonso García, E., et al. (2018). Towards global data products of Essential Biodiversity Variables on species traits. *Nature Ecology & Evolution* **2**, 1531–1540.

Koskela, J., Lefèvre, F., Schueler, S., Kraigher, H., Olrik, D. C., Hubert, J., Longauer, R., Bozzano, M., Yrjänä, L., Alizoti, P., Rotach, P., Vietto, L., Bordács, S., Myking, T., Eysteinsson, T., et al. (2013). Translating conservation genetics into management: pan-European minimum requirements for dynamic conservation units of forest tree genetic diversity. *Biological Conservation* **157**, 39–49.

*Kraemer, P. & Gerlach, G. (2017). Demerelate: calculating interindividual relatedness for kinship analysis based on codominant diploid genetic markers using R. *Molecular Ecology Resources* **17**, 1371–1377.

Laikre, L., Allendorf, F. W., Aroner, L. C., Baker, C. S., Gregovich, D. P., Hansen, M. M., Jackson, J. A., Kendall, K. C., McKelvey, K., Neel, M. C. & Olivieri, I. (2010). Neglect of genetic diversity in implementation of the convention on biological diversity. *Conservation Biology* **24**, 86–88.

Laikre, L., Hoban, S., Bruford, M. W., Segelbacher, G., Allendorf, F. W., Gajardo, G., Rodríguez, A. G., Hedrick, P. W., Heuertz, M., Hohenlohe, P. A., Jaffé, R., Johannesson, K., Liggins, L., MacDonald, A. J., et al. (2020). Post-2020 goals overlook genetic diversity. *Science* **367**, 1083–1085.

Laikre, L., Nilsson, T., Primmer, C. R., Ryman, N. & Allendorf, F. W. (2009). Importance of genetics in the interpretation of favourable conservation status. *Conservation Biology* **23**, 1378–1381.

Landguth, E. L., Cushman, S. A., Schwartz, M. K., McKelvey, K. S., Murphy, M. & Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology* **19**, 4179–4191.

Latombe, G., Pyšek, P., Jeschke, J. M., Blackburn, T. M., Bacher, S., Capinha, C., Costello, M. J., Fernández, M., Gregory, R. D., Hobern, D., Hui, C., Jetz, W., Kumschick, S., McGrannachan, C., Pergl, J., et al. (2017). A vision for global monitoring of biological invasions. *Biological Conservation* **213**, 295–308.

Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., Boehm, F., Caporaso, N. E., Cornelis, M. C., Edenberg, H. J., Gabriel, S. B., Harris, E. L., Hu, F. B., Jacobs, K. B., Kraft, P., et al. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic Epidemiology* **34**, 591–602.

Lawrence, E. R., Benavente, J. N., Matte, J.-M., Marin, K., Wells, Z. R. R., Bernos, T. A., Krasteva, N., Habrich, A., Nessel, G. A., Koumrouyan, R. A. & Fraser, D. J. (2019). Geo-referenced population-specific microsatellite data across American continents, the MacroPopGen Database. *Scientific Data* **6**, 14.

Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Segurel, L., Venkat, A., Andolfatto, P. & Przeworski, M. (2012). Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS Biology* **10**, e1001388.

Leigh, D. M., Hendry, A. P., Vázquez-Domínguez, E. & Friesen, V. L. (2019). Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. *Evolutionary Applications* **12**, 1505–1512.

Leigh, D. M., Van Rees, C. B., Millette, K. L., Breed, M. F., Schmidt, C., Bertola, L. D., Hand, B. K., Hunter, M. E., Jensen, E. L., Kershaw, F., Liggins, L., Luikart, G., Manel, S., Mergeay, J., Miller, J. M., et al. (2021). Opportunities and challenges of macrogenetic studies. *Nature Reviews Genetics* **22**, 791–807.

Leroy, G., Carroll, E. L., Bruford, M. W., DeWoody, J. A., Strand, A., Waits, L. & Wang, J. (2018). Next-generation metrics for monitoring genetic erosion within populations of conservation concern. *Evolutionary Applications* **11**, 1066–1083.

Liggins, L., Hudson, M. & Anderson, J. (2021a). Creating space for Indigenous perspectives on access and benefit-sharing: encouraging researcher use of the Local Contexts Notices. *Molecular Ecology* **30**, 2477–2482.

Liggins, L., Noble, C. & the Ira Moana Network (2021b). The Ira Moana Project: a genetic observatory for Aotearoa's marine biodiversity. *Frontiers in Marine Science* **8**, 740953.

*Lischer, H. E. L. & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298–229.

Lowe, A. J., Boshier, D., Ward, M., Bacles, C. F. E. & Navarro, C. (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity* **95**, 255–273.

Luikart, G., Antao, T., Hand, B. K., Muhlfeld, C. C., Boyer, M. C., Cosart, T., Trethewey, B., Al-Chockhachy, R. & Waples, R. S. (2021). Detecting population declines via monitoring the effective number of breeders (Nb). *Molecular Ecology Resources* **21**, 379–393.

Luikart, G., Kardos, M., Hand, B. K., Rajora, O. P., Aitken, S. N. & Hohenlohe, P. A. (2018). Population genomics: advancing understanding of nature. In *Population Genomics* (ed. O. Rajora), pp. 3–79. Springer, Cham.

Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K. & Allendorf, F. W. (2010). Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics* **11**, 355–373.

Maguire, K. C., Shinneman, D. J., Potter, K. M. & Hipkins, V. D. (2018). Intraspecific niche models for Ponderosa pine (*Pinus ponderosa*) suggest potential variability in population-level response to climate change. *Systematic Biology* **67**, 965–978.

*Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* **20**, 229–237.

Manel, S., Guerin, P.-E., Mouillot, D., Blanchet, S., Velez, L., Albouy, C. & Pellissier, L. (2020). Global determinants of freshwater and marine fish genetic diversity. *Nature Communications* **11**, 692.

*Marshall, T. C., Slate, J., Kruuk, L. E. & Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639–655.

Martins, K., Gugger, P. F., Llanderal-Mendoza, J., González-Rodríguez, A., Fitz-Gibbon, S. T., Zhao, J. L., Rodríguez-Correa, H., Oyama, K. & Sork, V. L. (2018). Landscape genomics provides evidence of

climate-associated genetic variation in Mexican populations of *Quercus rugosa*. *Evolutionary Applications* **11**, 1842–1858.
Mather, N., Traves, S. M. & Ho, S. Y. W. (2020). A practical introduction to sequentially Markovian coalescent methods for estimating demographic history from genomic data. *Ecology and Evolution* **10**, 579–589.
*Maunder, M., Lyte, B., Dransfield, J. & Baker, W. (2001). The conservation value of botanic garden palm collections. *Biological Conservation* **98**, 259–271.
McCartney, A., Anderson, J., Liggins, L., Hudson, M., Anderson, M., Te Aika, B., Geary, J., Cook-Deegan, R., Patel, H. R. & Phillippy, A. M. (2021). Balancing openness with Indigenous Data Sovereignty — an opportunity to leave no one behind in the journey to sequence all of life. Electronic file available at SSRN: https://ssrn.com/abstract=3922996. Accessed March 10, 2022.
*Meirmans, P. G. & Van Tienderen, P. H. (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* **4**, 792–794.
Mergeay, J. (2012). Afwegingskader voor de versterking van populaties van Europees beschermde soorten. Adviezen van het Instituut voor Natuur- en Bosonderzoek. INBO.A.2012.141.. Instituut voor Natuur- en Bosonderzoek, Brussel.
Meyer, C., Jetz, W., Guralnick, R. P., Fritz, S. A. & Kreft, H. (2016). Range geometry and socio-economics dominate species-level biases in occurrence information. *Global Ecology and Biogeography* **25**, 1181–1193.
Millette, K. L., Fugère, V., Debyser, C., Greiner, A., Chain, F. J. J. & Gonzalez, A. (2020). No consistent effects of humans on animal genetic diversity worldwide. *Ecology Letters* **23**, 55–67.
Mimura, M., Yahara, T., Faith, D. P., Vázquez-Domínguez, E., Colautti, R. I., Araki, H., Javadi, F., Núñez-Farfán, J., Mori, A. S., Zhou, S., Hollingsworth, P. M., Neaves, L. E., Fukano, Y., Smith, G. F., Sato, Y.-I., et al. (2017). Understanding and monitoring the consequences of human impacts on intraspecific variation. *Evolutionary Applications* **10**, 121–139.
*Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A. & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**, 1530–1534.
Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., Wang, Z., Rahbek, C., Marske, K. A. & Nogués-Bravo, D. (2016). An Anthropocene map of genetic diversity. *Science* **353**, 1532–1535.
Mohr, J.J., Harrison, P.A., Stanhope, J. & Breed, M.F. (2022). Is the genomics 'cart' before the restoration ecology 'horse'? Insights from qualitative interviews and trends from the literature. bioRxiv 2021.08.05.455206; https://doi.org/10.1101/2021.08.05.
Monteiro, W. P., Veiga, J. C., Silva, A. R., Carvalho, C. S., Lanes, É. C. M., Rico, Y. & Jaffé, R. (2019). Everything you always wanted to know about gene flow in tropical landscapes (but were afraid to ask). *PeerJ* **7**, e6446.
Muller-Karger, F. E., Miloslavich, P., Bax, N., Simmons, S., Costello, M. J., Sousa Pinto, I., Canonico, G., Turner, W., Gill, M., Montes, E., Best, B., Pearlman, J., Halpin, P., Dunn, D., Benson, A., et al. (2018). Advancing marine biological observations and data requirements of the complementary Essential Ocean Variables (EOVs) and Essential Biodiversity Variables (EBVs) frameworks. *Frontiers in Marine Science* **5**, 211.
*Muñoz-Fuentes, V., Green, A. J. & Negro, J. J. (2013). Genetic studies facilitated management decisions on the invasion of the ruddy duck in Europe. *Biological Invasions* **15**, 723–728.
Murphy, M. A., Evans, J. S., Cushman, S. A. & Storfer, A. (2008). Representing genetic variation as continuous surfaces: an approach for identifying spatial dependency in landscape genetic studies. *Ecography* **31**, 685–697.
*Mussmann, S. M., Douglas, M. R., Chafin, T. K., Douglas, M. E. & Jarman, S. (2019). BA3-SNPs: contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods in Ecology and Evolution* **10**, 1808–1813.
Nadachowska-Brzyska, K., Konczal, M. & Babik, W. (2022). Navigating the temporal continuum of effective population size. *Methods in Ecology and Evolution* **13**, 22–41.
Nater, A., Greminger, M. P., Arora, N., van Schaik, C. P., Goossens, B., Singleton, I., Verschoor, E. J., Warren, K. S. & Krützen, M. (2015). Reconstructing the demographic history of orang-utans using Approximate Bayesian Computation. *Molecular Ecology* **24**, 310–327.
Navarro, L. M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W. D., Londoño, M. C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M. J., Delavaud, A., El Serafy, G. Y., Ferrier, S., Geijzendorffer, I., Geller, G. N., et al. (2017). Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability* **29**, 158–169.
Nazareno, A. G., Bemmels, J. B., Dick, C. W. & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Mol Ecol Resour* **17**, 1136–1147. https://doi.org/10.1111/1755-0998.12654.
Neel, M. C., McKelvey, K., Ryman, N., Lloyd, M. W., Short Bull, R., Allendorf, F. W., Schwartz, M. K. & Waples, R. S. (2013). Estimation of effective population size in continuously distributed populations: there goes the neighborhood. *Heredity* **111**, 189–199.
Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10.
Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. K. & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in Genetics* **26**, 177–187.
Ovaskainen, O., Abrego, N., Somervuo, P., Palorinne, I., Hardwick, B., Pitkanen, J. M., Andrew, N. R., Niklaus, P. A., Schmidt, N. M., Seibold, S., Vogt, J., Zakharov, E. V., Hebert, P. D. N., Roslin, T. & Ivanova, N. V. (2020). Monitoring fungal communities with the Global Spore Sampling Project. *Frontiers in Ecology and Evolution* **7**, 511.
Pacioni, C., Wayne, A. F. & Page, M. (2019). Guidelines for genetic management in mammal translocation programs. *Biological Conservation* **237**, 105–113.
Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., Omrak, A., Vartanyan, S., Poinar, H., Götherström, A., Reich, D. & Dalén, L. (2015). Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. *Current Biology* **25**, 1395–1400.
*Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420.
Parsons, K. M., Everett, M., Dahlheim, M. & Park, L. (2018). Water, water everywhere: environmental DNA can unlock population structure in elusive marine species. *Royal Society Open Science* **5**, 180537.
Paz-Vinas, I., Jensen, E. L., Bertola, L. D., Breed, M. F., Hand, B. K., Hunter, M. E., Kershaw, F., Leigh, D. M., Luikart, G., Mergeay, J., Miller, J. M., van Rees, C. B., Segelbacher, G. & Hoban, S. (2021). Macrogenetic studies must not ignore limitations of genetic markers and scale. *Ecology Letters* **24**, 1282–1284.
Paz-Vinas, I., Loot, G., Hermoso, V., Veyssière, C., Poulet, N., Grenouillet, G. & Blanchet, S. (2018). Systematic conservation planning for intraspecific genetic diversity. *Proceedings of the Royal Society B: Biological Sciences* **285**, 20172746.
*Peakall, R. & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. *Bioinformatics* **28**, 2537–2539.
Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., Cardoso, A. C., Coops, N. C., Dulloo, E., Faith, D. P., Freyhof, J., Gregory, R. D., et al. (2013). Essential biodiversity variables. *Science* **339**, 277–278.
*Pew, J., Muir, P. H., Wang, J. & Frasier, T. R. (2015). related: an R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources* **15**, 557–561.
Pierson, J. C., Coates, D. J., Oostermeijer, J. G. B., Beissinger, S. R., Bragg, J. G., Sunnucks, P., Schumaker, N. H. & Young, A. G. (2016). Genetic factors in threatened species recovery plans on three continents. *Frontiers in Ecology and the Environment* **14**, 433–440.
Pinsky, M. L. & Palumbi, S. R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology* **23**, 29–39.
Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., Barrio-Amorós, C. L., Salazar-Valenzuela, D. & Prost, S. (2018). Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. *GigaScience* **7**, giy033.
Pope, L. C., Liggins, L., Keyse, J., Carvalho, S. B. & Riginos, C. (2015a). Not the time or the place: the missing spatio-temporal link in publicly available genetic data. *Molecular Ecology* **24**, 3802–3380.
Pope, L. C., Riginos, C., Ovenden, J., Keyse, J. & Blomberg, S. P. (2015b). Population genetic diversity in the Australian 'seascape': a bioregion approach. *PLoS One* **10**, e0136275.
Posledovich, D., Ekblom, R. & Laikre, L. (2021) *Mapping and Monitoring Genetic Diversity in Sweden: A Proposal for Species, Methods and Costs*. The Swedish Environmental Protection Agency Report 6959, ISBN 978-91-620-6959-9. Swedish Environmental Protection Agency, Naturvårdsverket, Sotckholm, Sweden.
Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.
Putman, A. I. & Carbone, I. (2014). Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution* **4**, 4399–4428.
Ralls, K., Ballou, J. D., Dudash, M. R., Eldridge, M. D., Fenster, C. B., Lacy, R. C., Sunnucks, P. & Frankham, R. (2018). Call for a paradigm shift in the genetic management of fragmented populations. *Conservation Letters* **11**, e12412.
Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S. J., Novella-Fernandez, R., Alberdi, A. & Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences* **116**, 10418–10423.

Razgour, O., Juste, J., Ibáñez, C., Kiefer, A., Rebelo, H., Puechmaille, S. J., Arlettaz, R., Burke, T., Dawson, D. A., Beaumont, M. & Jones, G. (2013). The shaping of genetic variation in edge-of-range populations under past and future climate change. *Ecology Letters* **16**, 1258–1266.

Razgour, O., Taggart, J. B., Manel, S., Juste, J., Ibanez, C., Rebelo, H., Alberdi, A., Jones, G. & Park, K. (2018). An integrated framework to identify wildlife populations under threat from climate change. *Molecular Ecology Resources* **18**, 18–31.

Reed, D. H. & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology* **17**, 230–237.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M. & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology* **24**, 4348–4370.

Rey, O., Eizaguirre, C., Angers, B., Baltazar-Soares, M., Sagonas, K., Prunier, J. G. & Blanchet, S. (2020). Linking epigenetics and biological conservation: towards a *conservation epigenetics* perspective. *Functional Ecology* **34**, 414–427.

Riginos, C., Crandall, E. D., Liggins, L., Gaither, M. R., Ewing, R. J., Meyer, C., Andrews, K., Euclide, P., Titus, B., Overgaard Therkildsen, N., Salces Castellano, A., Stewart, L. C., Toonen, R. J. & Deck, J. (2020). Building a global genomics observatory: using GEOME (the Genomic Observatories Metadatabase) to expedite and improve deposition and retrieval of genetic data and metadata for biodiversity research. *Molecular Ecology Resources* **20**, 1458–1469.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Dernat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J. M., Nabholz, B., Roux, C., Tsagkogeorga, G., et al. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* **515**, 261–263.

*Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M. A. & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 539–554.

*Rousset, F. (2008). Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* **8**, 103–106.

*Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E. & Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**, 3299–3302.

Ryman, N., Laikre, L. & Hössjer, O. (2019). Do estimates of contemporary effective population size tell us what we want to know? *Molecular Ecology* **28**, 1904–1918.

Schlötterer, C. (2004). The evolution of molecular markers — just a matter of fashion? *Nature Reviews Genetics* **5**, 63–69.

Schmeller, D. S., Julliard, R., Bellingham, P. J., Böhm, M., Brummitt, N., Chiarucci, A., Couvet, D., Elmendorf, S., Forsyth, D. M., Garcia Moreno, J., Gregory, R. D., Magnusson, W. E., Martin, L. J., McGeoch, M. A., Mihoub, J. B., et al. (2015). Towards a global terrestrial species monitoring program. *Journal for Nature Conservation* **25**, 51–57.

Schmeller, D. S., Mihoub, J. B., Bowser, A., Arvanitidis, C., Costello, M. J., Fernandez, M. J., Geller, G. N., Hobern, D., Kissling, W. D., Regan, E. C., Saarenmaa, H., Turak, E. & Isaac, N. J. B. (2017). An operational definition of essential biodiversity variables. *Biodiversity and Conservation* **26**, 2967–2972.

Schmeller, D. S., Weatherdon, L. V., Loyau, A., Bondeau, A., Brotons, L., Brummitt, N., Geijzendorffer, I. R., Haase, P., Kuemmerlen, M., Martin, C. S., Mihoub, J. B., Rocchini, D., Saarenmaa, H., Stoll, S. & Regan, E. C. (2018). A suite of essential biodiversity variables for detecting critical biodiversity change. *Biological Reviews* **93**, 55–71.

Schmidt, C., Domaratzki, M., Kinnunen, R. P., Bowman, J. & Garroway, C. J. (2020). Continent-wide effects of urbanization on bird and mammal genetic diversity. *Proceedings of the Royal Society B* **287**, 20192497.

Schwartz, M. K., Luikart, G. & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution* **22**, 25–33.

*Schwartz, M. K. & McKelvey, K. S. (2009). Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics* **10**, 441–452.

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W. & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution* **8**, 907–917.

Sherwin, W. B., Chao, A., Jost, L. & Smouse, P. E. (2017). Information theory broadens the spectrum of molecular ecology and evolution. *Trends in Ecology & Evolution* **32**, 948–963.

Shinneman, D. J., Means, R. E., Potter, K. M. & Hipkins, V. D. (2016). Exploring climate niches of ponderosa pine (*Pinus ponderosa* Douglas ex Lawson) haplotypes in the western United States: implications for evolutionary history and conservation. *PLoS One* **11**, e0151811.

Shirk, A. J., Landguth, E. L. & Cushman, S. A. (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Molecular Ecology Resources* **17**, 1308–1317.

Sibbett, B., Rieseberg, L. H. & Narum, S. (2020). The Genomic Observatories Metadatabase. *Molecular Ecology Resources* **20**, 1453–1454.

Silva, A. R., Resende-Moreira, L. C., Carvalho, C. S., Lanes, E. C. M., Ortiz-Vera, M. P., Viana, P. L. & Jaffé, R. (2020). Range-wide neutral and adaptive genetic structure of an endemic herb from Amazonian Savannas. *AoB PLANTS* **12**, plaa003.

Sjödin, P., Kaj, I., Krone, S., Lascoux, M. & Nordborg, M. (2005). On the meaning and existence of an effective population size. *Genetics* **169**, 1061–1070.

*Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.

*Swofford, D. L. (2003). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer Associates, Sunderland.

*Szpiech, Z. A., Jakobsson, M. & Rosenberg, N. A. (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* **21**, 2498–2504.

Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P. A., Bona, E., Bovio, M. & Choler, P. (2012). Genetic diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters* **15**, 1439–1448.

Tahsin, T., Weissenbacher, D., Rivera, R., Beard, R., Firago, M., Wallstrom, G., Scotch, M. & Gonzalez, G. (2016). A high-precision rule-based extraction system for expanding geospatial metadata in GenBank records. *Journal of the American Medical Informatics Association* **23**, 934–941.

Tallmon, D. A., Gregovich, D., Waples, R. S., Baker, S. C., Jackson, J., Taylor, B. L., Archer, E., Martien, K. K., Allendorf, F. W. & Schwartz, M. K. (2010). When are genetic methods useful for estimating contemporary abundance and detecting population trends? *Molecular Ecology Resources* **10**, 684–692.

*Tallmon, D. A., Koyuk, A., Luikart, G. & Beaumont, M. A. (2008). ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources* **8**, 299–230.

Tenesa, A., Navarro, P., Hayes, B. J., Duffy, D. L., Clarke, G. M., Goddard, M. E. & Visscher, P. M. (2007). Recent human effective population size estimated from linkage disequilibrium. *Genome Research* **17**, 520–526.

Theodoridis, S., Fordham, D. A., Brown, S. C., Li, S., Rahbek, C. & Nogues-Bravo, D. (2020). Evolutionary history and past climate change shape the distribution of genetic diversity in terrestrial mammals. *Nature Communications* **11**, 2557.

Thomson, A., Archer, F., Coleman, M. A., Gajardo, G., Goodall-Copestake, W., Hoban, S., Laikre, L., Miller, A., O'Brien, D., Pérez-Espona, S., Segelbacher, G., Serrao, E., Sjøtun, K. & Stanley, M. (2021). The missing piece: genetic diversity in the UN decade of ocean science. *Evolutionary Applications* **14**, 1497–1518.

Toczydlowski, R., Liggins, L., Gaither, M. R., Anderson, T. J., Barton, R. L., Berg, J. T., Beskid, S. G., Davis, B., Delgado, A., Farrell, E., Ghoojaei, M., Himmelsbach, N., Holmes, A. E., Queeno, S. R., Trinh, T., et al. (2021). Poor data stewardship will hinder global genetic diversity surveillance. *PNAS* **118**, e2107934118.

Trivedi, C., Cavers, S., Atkinson, N., Clark, J. & Cottrell, J. (2019). *A Strategy for UK Forest Genetic Resources, Royal Botanic Gardens, Kew*. Board of Trustees of the Royal Botanic Gardens, Kew.

Turak, E., Harrison, I., Dudgeon, D., Abell, R., Bush, A., Darwall, W., Finlayson, C. M., Ferrier, S., Freyhof, J., Hermoso, V., Juffe-Bignoli, D., Linke, S., Nel, J., Patricio, H. C., Pittock, J., et al. (2017). Essential biodiversity variables for measuring change in global freshwater biodiversity. *Biological Conservation* **213**, 272–279.

Turon, X., Antich, A., Palacín, C., Præbel, K. & Wangensteen, O. S. (2020). From metabarcoding to metaphylogeography: separating the wheat from the chaff. *Ecological Applications* **30**, e02036.

Valentini, R., Marchesini, L. B., Gianelle, D., Sala, G., Yarovslavtsev, A., Vasenev, V. & Castaldi, S. (2019). New tree monitoring systems: from Industry 4.0 to Nature 4.0. *Annals of Silvicultural Research* **43**, 84–88.

Van Oppen, M. J., Gates, R. D., Blackall, L. L., Cantin, N., Chakravarti, L. J., Chan, W. Y., Cormick, C., Crean, A., Damjanovic, K., Epstein, H., Harrison, P. L., Jones, T. A., Miller, M., Pears, R. J. & Peplow, L. M. (2017). Shifting paradigms in restoration of the world's coral reefs. *Global Change Biology* **23**, 3437–3448.

Vandergast, A. G., Perry, W. M., Lugo, R. V. & Hathaway, S. A. (2011). Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. *Molecular Ecology Resources* **11**, 158–161.

*Vangestel, C., Mergeay, J., Dawson, D. A., Callens, T., Vandomme, V. & Lens, L. (2012). Genetic diversity and population structure in contemporary house sparrow populations along an urbanization gradient. *Heredity* **109**, 163–172.
*Verity, R. & Nichols, R. A. (2016). Estimating the number of subpopulations (K) in structured populations. *Genetics* **203**, 1827–1183.
von Takach, B., Penton, C. E., Murphy, B. P., Radford, I. J., Davies, H. F., Mill, B. M. & Banks, S. C. (2021). Population genomics and conservation management of a declining tropical rodent. *Heredity* **126**, 763–775.
Vranckx, G., Jacquemyn, H., Muys, B. & Honnay, O. (2012). Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology* **26**, 228–237.
Vranken, S., Wernberg, T., Scheben, A., Severn-Ellis, A., Batley, J., Bayer, P. E., Edwards, D., Wheeler, D. & Coleman, M. A. (2021). Genotype-environment mismatch of kelp forests under climate change. *Molecular Ecology* **30**, 3730–3746.
*Wang, J. (2011). COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources* **11**, 141–114.
Waples, R. S. (2016). Tiny estimates of the N e/N ratio in marine fishes: are they real? *Journal of Fish Biology* **89**, 2479–2504.
Waples, R. S. (2021). Relative precision of the sibship and LD methods for estimating effective population size with genomics-scale datasets. *Journal of Heredity* **112**, 535–539.
Waples, R. S. & England, P. R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics* **189**, 633–644.
Waples, R. S., Hoelzel, R. A., Gaggiotti, O., Tiedemann, R., Palsbøll, P. J., Cipriano, F., Jackson, J., Bickham, J. W. & Lang, A. R. (2018). Guidelines for genetic data analysis. *Journal of Cetacean Research and Management* **18**, 33–80.
Waples, R. S. & Lindley, S. T. (2018). Genomics and conservation units: the genetic basis of adult migration timing in Pacific salmonids. *Evolutionary Applications* **11**, 1518–1526.
Waples, R. S., Nammack, M., Cochrane, J. F. & Hutchings, J. A. (2013). A tale of two acts: endangered species listing practices in Canada and the United States. *BioScience* **63**, 723–734.
Wasserman, T. N., Cushman, S. A., Schwartz, M. K. & Wallin, D. O. (2010). Spatial scaling and multi-model inference in landscape genetics: *Martes americana* in northern Idaho. *Landscape Ecology* **25**, 1601–1612.
Wasserman, T. N., Cushman, S. A., Shirk, A. S., Landguth, E. L. & Littell, J. S. (2012). Simulating the effects of climate change on population connectivity of American marten (*Martes americana*) in the northern Rocky Mountains, USA. *Landscape Ecology* **27**, 211–225.
Wheat, R. E., Allen, J. M., Miller, S. D. L., Wilmers, C. C. & Levi, T. (2016). Environmental DNA from residual saliva for efficient noninvasive genetic monitoring of brown bears (*Ursus arctos*). *PLoS One* **11**, e0165259.
Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C. & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in Ecology & Evolution* **30**, 42–49.
Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T. & Vieglais, D. (2012). Darwin Core: an evolving community-developed biodiversity data standard. *PLoS One* **7**, e29715.
Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018.
*Willoughby, J. R., Sundaram, M., Wijayawardena, B. K., Kimble, S. J. A., Ji, Y., Fernandez, N. B., Antonides, J. D., Lamb, M. C., Marra, N. J. & DeWoody, J. A. (2015). The reduction of genetic diversity in threatened vertebrates and new recommendations regarding IUCN conservation rankings. *Biological Conservation* **191**, 495–503.
Winn, A. A., Elle, E., Kalisz, S., Cheptou, P.-O., Eckert, C. G., Goodwillie, C., Johnston, M. O., Moeller, D. A., Ree, R. H., Sargent, R. D. & Vallejo-Marín, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution* **65**, 3339–3359.
*Winter, D. J. (2012). MMOD: an R library for the calculation of population differentiation statistics. *Molecular Ecology Resources* **12**, 1158–1160.
Wooley, J. C., Field, D. & Glöckner, F. (2009). Extending standards for genomics and metagenomics data: a Research Coordination Network for the Genomic Standards Consortium (RCN4GSC). *Standards in Genomic Sciences* **1**, 87–90.
Wright, B. R., Grueber, C. E., Lott, M., Johnson, R. N. & Hogg, C. J. (2019). Impact of reduced-representation sequencing protocols on detecting population structure in a threatened marsupial. *Molecular Biology Reports* **46**, 5575–5580.
Wright, S. (1969). *Evolution and the Genetics of Populations* (Volume **2**). University of Chicago Press, Chicago.
Wright, S. (1978). *Evolution and the Genetics of Populations* (Volume **4**). University of Chicago Press, Chicago.
Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., Gilbert, J. A., Karsch-Mizrachi, I., Johnston, A., Cochrane, G., Vaughan, R., Hunter, C., Park, J., Morriso, N., Rocca-Serra, P., et al. (2011). Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology* **29**, 415–420.

IX. Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Workflow for genetic data processing with examples of common, standard file formats and software used to manipulate and edit files.
Table S1. Examples of common, standard analysis methods and software to measure Essential Biodiversity Variables (EBVs).
Appendix S1. Additional discussion on ranking the scalability of EBVs.
Appendix S2. Consideration of Hybridization as an EBV.
Appendix S3. Obtaining genetic data.
Appendix S4. Brief discussion of indicators.

(*Received 23 October 2020; revised 25 February 2022; accepted 2 March 2022; published online 12 April 2022*)

