Annual Review of Genetics

Museum Genomics

Daren C. Card,[1,2] Beth Shapiro,[3,4] Gonzalo Giribet,[1,2] Craig Moritz,[5] and Scott V. Edwards[1,2]

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA; email: sedwards@fas.harvard.edu

[2]Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA

[3]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California 95064, USA

[4]Howard Hughes Medical Institute, University of California, Santa Cruz, California 95064, USA

[5]Centre for Biodiversity Analysis and Research School of Biology, The Australian National University, Canberra, Australian Capital Territory 0200, Australia

Annu. Rev. Genet. 2021. 55:633–59

First published as a Review in Advance on September 23, 2021

The *Annual Review of Genetics* is online at genet.annualreviews.org

https://doi.org/10.1146/annurev-genet-071719-020506

Keywords

Anthropocene, ancient DNA, cryogenic collections, museomics, museum curation, natural history collections

Abstract

Natural history collections are invaluable repositories of biological information that provide an unrivaled record of Earth's biodiversity. Museum genomics—genomics research using traditional museum and cryogenic collections and the infrastructure supporting these investigations—has particularly enhanced research in ecology and evolutionary biology, the study of extinct organisms, and the impact of anthropogenic activity on biodiversity. However, leveraging genomics in biological collections has exposed challenges, such as digitizing, integrating, and sharing collections data; updating practices to ensure broadly optimal data extraction from existing and new collections; and modernizing collections practices, infrastructure, and policies to ensure fair, sustainable, and genomically manifold uses of museum collections by increasingly diverse stakeholders. Museum genomics collections are poised to address these challenges and, with increasingly sensitive genomics approaches, will catalyze a future era of reproducibility, innovation, and insight made possible through integrating museum and genome sciences.

1. INTRODUCTION

Natural history collection: a collection of preserved specimens or specimen-derived parts that is curated with the purpose of illuminating organismic natural history

Voucher specimen: a representative sample retained as a reference for an identified taxon and permanently deposited into an accessible collection facility

Metadata: a set of data that provides information about other data, namely the specimens that form biological collections

Natural history collections are essential repositories of biological information and have become the foundation for diverse fields of research generating significant knowledge of the natural world over the past several centuries (105). Natural history collections harbor deep, taxon-specific information derived from global populations of plants, animals, fungi, and microorganisms, and the specimens in natural history collections provide the foundation for diverse areas of basic science and research of relevance to human health and pressing societal issues, including zoonotic pathogens and climate change (72, 76, 79, 82, 87, 123). These collections vary significantly in form and function due to fundamental differences in the characteristics of specimens or associated research programs, shifts in curatorial practices and funding over time, and ever-changing technological advancements that increase access to biological information. Due to rapid developments in high-throughput sequencing, natural history specimens amassed over the last several centuries—often including specimens of long-extinct species and populations—are now accessible to genomic analysis (6, 22). Additionally, over the last 50 years, natural history museums have systematically collected and cryogenically stored field-collected tissues and other genomic resources that fuel diverse research in comparative biology, population genetics, and genomics. These behind-the-scenes collections may not be visually stunning or accessible to the public; therefore, many laypersons and scientists are unaware of their existence. Nonetheless, both traditional museum and genomically archived specimens in biological collections now form a critical infrastructure for many facets of modern biology and research that serve society (30). Diverse types of institutions, from zoological museums and herbaria to arboreta, living-stock collections, zoos, and aquaria, all contribute to this scientific infrastructure (105). This review focuses on the contributions of zoological specimens and cryogenic collections to modern genomics and comparative biology and outlines challenges and opportunities facing these important resources.

2. NATURAL HISTORY COLLECTIONS HARBOR UNRIVALED GENOMIC INFORMATION

2.1. Building Traditional Natural History Collections

The international program of biodiversity preservation pursued by natural history museums is vast and ambitious, amounting to nothing less than the documentation of every species—and often multiple populations per species—and their associated environments. Traditional natural history collections contain diverse materials and have historically focused on preserving voucher specimens as (*a*) dried specimens such as skins of birds and mammals, fossils, pinned insects, and shells of molluscs and other invertebrates and (*b*) specimens stored in liquid preservatives such as alcohol or formalin, an approach often taken to preserve fishes, amphibians, reptiles, and invertebrates. Many of these specimens have been collected by field biologists visiting particular geographic regions and either sampling the local community of organisms largely randomly (i.e., general collecting) or targeting certain taxonomic groups for specific research purposes. Diverse analyses of genomics, morphology, chemistry, and specimen metadata have therefore often been applied to specimens collected decades earlier and for reasons completely unrelated to their current use in research (68, 143, 152).

2.2. Modern Cryogenic Collection Practices

Many museums have also invested in housing biological tissue samples collected specifically for a range of molecular investigations—so-called tissue collections or, more elegantly, collections of genetic/genomic resources, which have grown dramatically in recent decades (**Figure 1**). Curatorial procedures for tissue samples vary greatly and have continually changed over time as

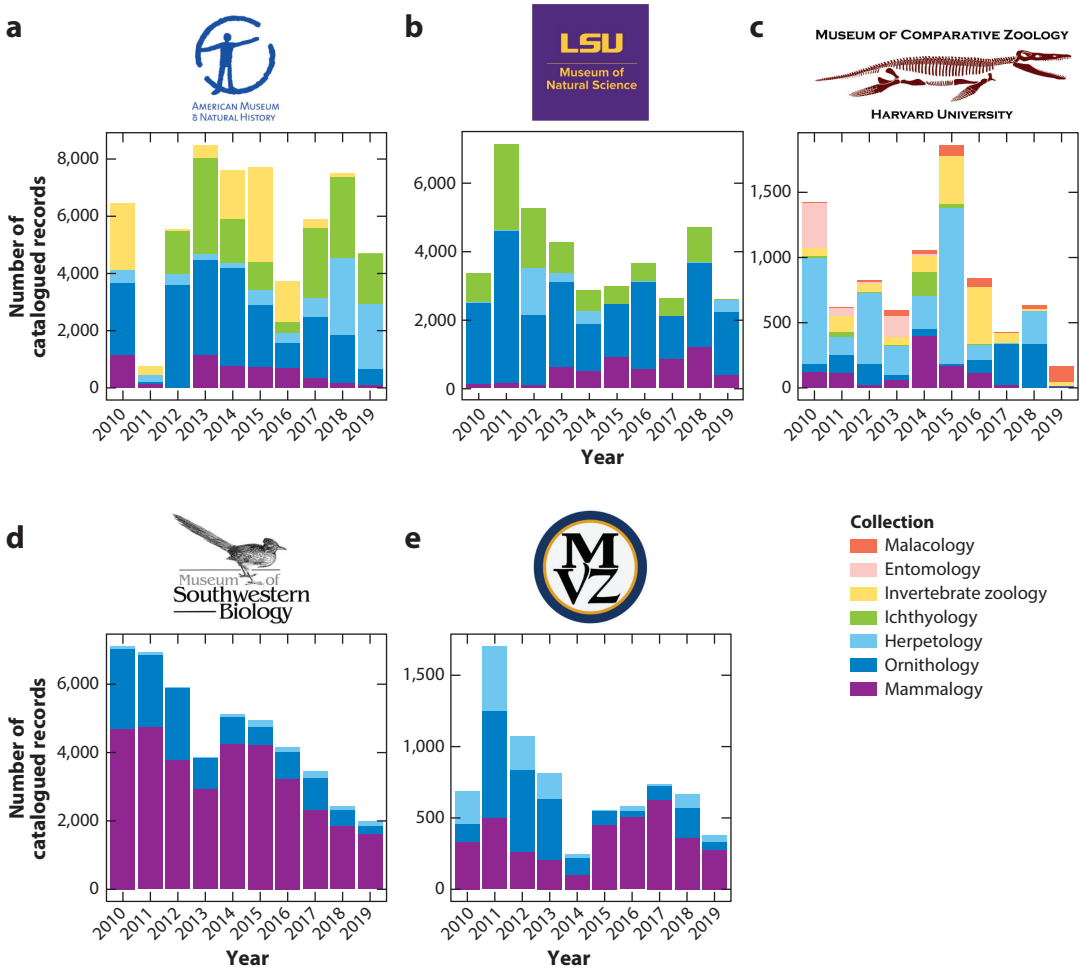


Figure 1

Temporal patterns of cryogenic collection growth at five major museum natural history collections. Annual catalogued records per collection for (*a*) American Museum of Natural History (AMNH), (*b*) Louisiana State University Museum of Natural Science (LSUMZ), (*c*) Harvard University Museum of Comparative Zoology (MCZ), (*d*) University of New Mexico Museum of Southwestern Biology (MSB), and (*e*) University of California Museum of Vertebrate Zoology (MVZ). All counts only reflect newly collected and catalogued samples with an associated voucher specimen, but each of these collections also has significant numbers of samples with no associated voucher specimens and previously collected samples that have been more recently added to cryogenic collections. For example, MSB has collected and cryopreserved 8,886 fish genomic samples between 2010 and 2019, but none had associated voucher specimens. Raw cryogenic collection count data is available online at **https://doi.org/10.5281/zenodo.5093840** and **https://edwards-bird-lab.github.io/museum-genomics/**. See **Supplemental Figure 1** for cumulative sums of catalogued records per collection over this time period.

Supplemental Material >

the demands of molecular techniques have evolved. Some genomic resource collections are stored at room temperature either without preservatives or in simple buffers, such as ethanol (104). In many other collections, tissues are stored at cooler temperatures, such as 4°C for samples not intended for freezing, or at bio-stable temperatures, such as in −80°C freezers or vapor liquid nitrogen cryogenic collections (**Figure 2**). Paleontological collections have also begun to store fossils at either cold or frozen temperatures to limit the further decay of potentially preserved

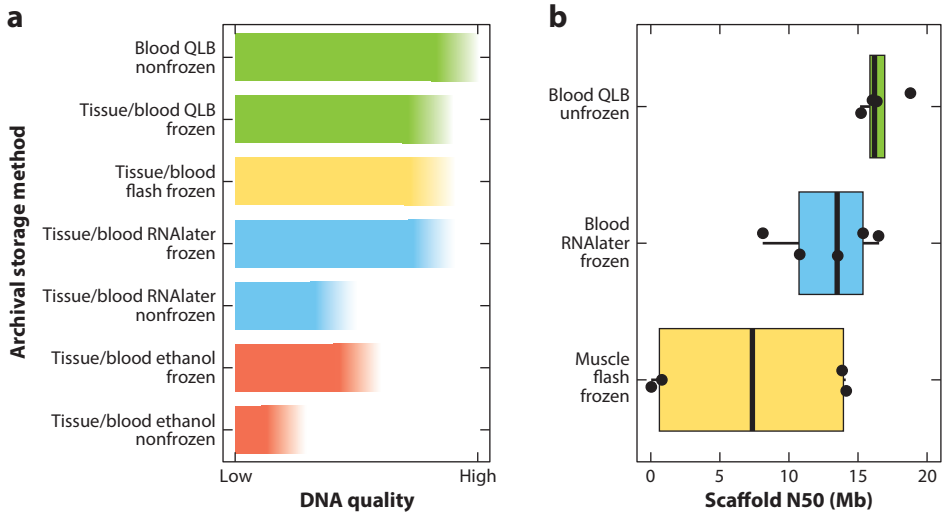


Figure 2

Museum preservation considerations for genomics approaches. (*a*) Hypothetical variation in long-term potential for genomics investigations for common genomics archival storage methods. Potential is based on DNA quality, and certain archival storage methods are known to perform poorly in downstream transcriptomic and epigenomic investigations, such as material stored in Queen's lysis buffer (QLB) or ethanol. (*b*) Empirical patterns of genome assembly scaffold contiguity (N50) for avian genome assemblies sequenced in the lab of S.V. Edwards using 10x Genomics linked-read technology based on tissue type and preservation protocol. Muscle samples were flash frozen and stored at −80°C. Blood samples in RNAlater were frozen in liquid nitrogen within one day and were stored at −80°C. Blood samples in QLB were kept at ambient temperatures for one week to one month before medium-term or permanent storage at 4°C. Sequencing coverage varied between approximately 28× and 70× and had no apparent impact on assembly contiguity. Raw genome assembly contiguity data is available online at **https://doi.org/10.5281/zenodo.5093840** and **https://edwards-bird-lab.github.io/museum-genomics/**.

genetic materials (114, 117). The utility of such collections for future research is directly related to the care in collecting and preserving tissue samples, which highlights the need to investigate and apply new, standardized curatorial procedures that maximally preserve biological information (134).

Researchers have deposited a range of materials into genomic collections, including small tissue samples, most often including liver, heart, or muscle samples for vertebrates and some invertebrates, and whole organisms or nonspecific tissue samples for most invertebrates. Increasingly, genomic resource collections have diversified to include additional tissue types and samples preserved with the care required for more specialized genomic techniques, such as whole-genome sequencing or transcriptomics (**Figure 2**). Additionally, environmental samples, such as sediment and water samples and environmental DNA (eDNA) filters, as well as specimens such as nests, middens, and intestinal tracts for microbiome studies, are collected and stored for metagenomic and eDNA analyses (45, 100). Cryospecimens are usually closely tied to the physical specimens from which they came or, for many smaller specimens such as invertebrates, associated with other individuals from the same collecting effort. This coarchiving of genomic resources and associated specimens allows a level of replication and verification that is unusual in biological science, enabling the possibility of revisiting individual specimens for verification or re-evaluation of hypotheses (72, 123, 134). However, many collections today still bear the imprint of collecting practices used in the past, strongly valuing the preservation of phenotypes for anatomical and

Environmental DNA (eDNA): DNA that is collected from a variety of environmental samples rather than sampled directly from individual organisms

Metagenomics: the study of collections of genomic samples obtained from mixed communities of organisms, such as microbial populations in the gut

imaging work, and are only recently expanding to incorporate collection and preservation methods to serve the broader genomics community. Studies attempting to compare the preservation of macromolecules under different preservation regimes are few and necessarily small in temporal scope (28, 102). Numerous protocols for preserving frozen tissues for DNA analysis are available (75, 77, 116), and best practices are shared throughout the community through groups such as the Society for the Preservation of Natural History Collections; increasingly, protocols for preserving traditional specimens and fossils for the eventual extraction of DNA and other macromolecules have appeared (24, 56, 70, 149).

2.3. Considerations for Preserving Genomic Resources

Challenges of preservation have been apparent throughout the history of DNA recovery from preserved specimens. The first wave of tissue collections growth occurred in the 1970s, when the prevalent method for genetic analysis was allozyme electrophoresis. Allozymes, or protein variants, were visualized by staining whole animal extracts run through a starch gel and were often retrievable from tissues sampled many hours after animal sacrifice in the field. Later, Sanger sequencing technology replaced allozyme analysis as the standard approach for comparative genomic analysis, first via direct sequencing of RNA (e.g., 47) and soon after via amplified DNA using polymerase chain reaction (PCR). Because PCR can amplify from extremely limited amounts of DNA and target relatively short fragment lengths, all but the poorest-quality tissue samples (excluding samples in paleontological collections) are usually sufficient for extracting useful molecular information, especially from high copy number loci like those found in mitochondria or the nuclear ribosomal RNAs (80, 97). Ancient DNA investigations had their own unique challenges that limited research, including contamination and short and damaged DNA molecules from the processes of cross-linking and cytosine deamination (22, 113), which significantly complicated most PCR investigations. The vast majority of museum tissue collections were initiated during the decades when allozymes and, especially, PCR were the workhorse technologies of evolutionary genomics. This contingency has cast a long shadow on the place and utility of tissue collections in today's era of high-throughput sequencing. For example, few tissue collections in natural history museums today are amenable to analysis of RNA, which decays much more rapidly than DNA postmortem; some RNA-ready collections are currently being assembled, but they are still rare.

The ongoing long-read sequencing revolution, with average and maximum read lengths exceeding 10 kb and 1 Mb, respectively, has abruptly constrained the number of usable existing genomic samples in museum collections, although some studies report improved results on museum tissues compared to short-read methods (15, 32). Moreover, the proliferation of functional genomics assays that rely on unstable RNA molecules or freshly extracted cells—techniques such as transcriptomics, epigenomics, and methylomics—often require fresh or carefully preserved material that is largely absent from natural history collections, though some institutions have prioritized flash-frozen collections since the 1970s (41). Today, museum tissue collections are in a dramatic process of recollecting, evaluating practices, and determining new protocols to best serve the study of the genomics of biodiversity (43).

3. MUSEUM GENOMICS DRIVES DIVERSE RESEARCH AREAS OF RELEVANCE TO BIODIVERSITY AND SOCIETY

3.1. Evolutionary Investigations in Museum Genomics

Biodiversity science has long been a primary field of inquiry in natural history museum–based research, a trend that has continued with today's genomics-fueled museum science programs.

Society for the Preservation of Natural History Collections: an international organization devoted to the preservation, conservation, and management of natural history collections

Ancient DNA: DNA successfully isolated from degraded specimens; the field that studies these materials is often also called ancient DNA

Cross-linking: formalin treatment is used to fix tissues, inducing covalent bonds within and between DNA and proteins

Cytosine deamination: the postmortem accumulation of uracil at cytosine nucleotides induced by spontaneous loss of the amine group found in cytosine

Epigenomics: the study of genomic chemical modifications that induce changes in gene expression that cannot be explained by DNA sequence modifications

Methylomics: the genomic investigation of nucleic acid methylation modifications in genomes, which can vary between tissues, individuals, or species

Genomics-engaged biodiversity science focuses on discovering, cataloging, and understanding relationships and interactions between populations of living organisms, encompassing numerous subfields—such as phylogenetics, phylogeography, and population genetics—of the broader discipline of evolutionary biology. Recently, museum collections have driven genomic investigations in comparative biology in diverse clades across the tree of life (e.g., 17, 27, 46, 63, 74). Species-specific investigations have also targeted deeper sampling of wild populations to detect microevolutionary patterns and adaptive evolution. However, such studies have their greatest impact and allow the possibility of replication and verification when genomic data are integrated with phenotypic data derived from museum specimens. North American house mice exhibit latitudinal variation in several traits, including body size (i.e., Bergmann's rule), suggesting adaptive evolution. Leveraging these patterns, population sampling across the cline, and genetic crosses in common garden conditions, investigators have identified genetic variation associated with differential gene expression—known as expression quantitative trait loci (eQTLs)—that is also correlated with latitude, implicating numerous genes as potential drivers of adaptive phenotypic evolution (90, 112). In contrast to many similar studies that fail to archive the specimens used to evaluate phenotypes, these researchers deposited all wild-caught samples and captive progeny into the University of California, Berkeley, Museum of Vertebrate Zoology collections. Doing so facilitates future genomics investigations and allows re-evaluation and verification of scientific claims.

Researchers have also leveraged the repeated evolution of life-history characteristics across disparate lineages as a means of understanding the process of convergent evolution at both the organismal level and the molecular level (84). Sackton et al. (133) used museum holdings to generate genome assemblies for 11 paleognathous birds and used these resources to identify conserved noncoding regions that were likely regulatory and had undergone lineage-specific accelerations, suggesting loss or change of function across one or more flightless species. Epigenetic approaches, namely ATAC-seq (assay for transposase-accessible chromatin with high-throughput sequencing) (20), improved functional understanding of these genomes, although material for these downstream studies was not available in museum collections and had to be newly collected. This study exemplifies one that straddled two worlds: using existing museum tissue collections for comparative genomics but requiring additional collections to profile genome-wide epigenetic states and gene expression. Finally, studies of genome evolution itself—previously the domain of geneticists without museum training—have leveraged the rich holdings of museum genomics collections to profile molecular phenotypes that were previously inaccessible to museum investigators. For instance, Pasquesi et al. (111) profiled repetitive elements across 66 squamate reptile species derived mostly from museum genomic samples and found marked differences in repeat abundance and composition, despite very similar overall genome sizes, which has challenged existing hypotheses on the evolution of genome size based on investigations in mammals and birds. Indeed, as genomics further ingrains itself in museum science and evolves, we expect that increasingly diverse molecular phenotypes will be measured directly from museum specimens, enabling new connections between genotype and phenotype and diverse investigations of comparative biology.

3.2. Ecological Insights from Museum Genomics

Museum cryogenic collections have also facilitated decades of study in the field of ecology. Because museum specimens typically have precise sampling locality and timing data, studies have leveraged genomics to understand how natural populations of organisms respond to strong ecological pressures. For example, Friis et al. (52) used a museum-derived reference genome and

reduced representation genome sequencing of Oregon juncos (*Junco hyemalis oreganus*) to explore correlations between genomic and environmental variance, which yielded evidence of strong drift in isolated populations and rapid, environment-driven local adaptation between populations with no obvious geographic barriers to gene flow. Parejo et al. (110) sequenced whole genomes of 22 historic honey bee samples collected between 1879 and 1959 from Switzerland and housed in the Natural History Museum in Bern. Genetic comparisons with modern samples from the same region found that genetic diversity was similar between time periods, which may be due to modern apicultural practices. Moreover, the investigators identified signals of selection associated with genes linked to xenobiotic response physiology between historic and modern bee populations, which suggests that agricultural and apicultural chemical usage over the last century may have strong ecological consequences on these important pollinators. A similar study utilizing dozens of modern and historic samples from two butterfly species deposited in the Finnish Museum of Natural History instead found evidence for loss of genetic diversity in both species since the early 1900s, likely due to the impacts of human activity during this time period (53).

Genomics approaches can be used to validate putative hybrids between two distinct species or uncover otherwise hidden relationships between organisms due to relatedness in pedigree or parentage analyses. Toews et al. (148) leveraged museum collections of parental species and genomics approaches to opportunistically detect an intergeneric hybrid between a blue-winged warbler (*Vermivora cyanoptera*) and a cerulean warbler (*Setophaga cerulea*). Yang et al. (159) utilized genomic samples from museums to design a genotype panel capable of discerning familial relationships between bison, which enables investigations of free-ranging and agricultural populations of this species. The increased coarchiving of parasites and their vertebrate and invertebrate hosts has supported numerous insights into the ecology and coevolution of parasite and host communities (11, 115). Moreover, armed with revolutionary knowledge of the role that microorganisms play in organismal biology from investigations of human health and disease, museum practitioners have driven metagenomic investigations of the microbiome in nonmodel species sampled from museum collections of mammals (18) and fishes (66).

3.3. Museum Genomics and Disease

Museum genomics collections have also contributed significantly to our understanding of infectious diseases in wild species and their geographic and temporal patterns of spread (44, 135). In one of the best-known uses of museum collections to identify the origin and prevalence of pathogens infectious to humans, rodent specimens and archived tissues in the Museum of Southwestern Biology at the University of New Mexico allowed researchers to confirm the hantavirus as the causative agent of pulmonary illnesses and deaths in the Four Corners region of the United States in the early 1990s (160). Likewise, museum specimens and archived tissues have played an important role in our understanding of the origin and spread of the devastating amphibian fungal pathogen chytrid *Batrachochytrium dendrobatidis* (Bd) (48, 124). Moreover, Hydeman et al. (73) and Carvalho et al. (26) used quantitative PCR (qPCR) to estimate the prevalence of Bd in frogs of Central African islands and to verify the Bd infection status of Brazilian frogs as assessed by visual inspection of tadpole mouthparts, respectively. However, the utility of museum collections for informing responses to emerging pathogens has been underestimated, and often, reference collections of potential reservoir species for pathogens are not made. Collections can provide epidemiologists with rapid genomic understanding of reservoir species and the emergence and evolutionary spread of a pathogen within and between species (31, 42). Overall, genomics technologies are powerful tools for illuminating organismal ecology and will be invaluable for future investigations of the natural and anthropogenic drivers of biological change.

4. MUSEUM GENOMICS PROVIDES A WINDOW ON THE ANTHROPOCENE AND ITS IMPACTS ON BIODIVERSITY

4.1. Early Insights into Human-Mediated Global Change

The past century has witnessed accelerating human-driven change and disruption of natural systems, with mostly dire consequences for biodiversity and ecosystem services. This unfolding sixth mass extinction (7), dubbed the Anthropocene (38), has resulted in an unplanned experiment providing insight into the resilience of species and ecosystems to rapid change and the capacity for rapid and sustainable evolution (101). The application of strong, human-mediated selection pressures—such as widespread chemical control of weeds, insects, and invasive mammals—has generated classic examples of the rapid evolution of resistance, with historical collections often serving as benchmarks (81, 96). By comparing the exomes of invasive rabbits collected before and after the myxomatosis pandemic that struck Europe and Australia beginning in the 1950s, Alves et al. (3) demonstrated that resistance to this biological control agent was polygenic and evolved in parallel in independent populations of rabbits from standing variation. Coalescent analyses of diversity across whole genomes have also shed new light on the historical changes in population size and efficiency of selection in species recently negatively impacted by humans, such as the extinct passenger pigeon (103) and several North American bird species (144). A recent study of eight extinct Australian rodent species, however, found no evidence of reduced genetic diversity in museum specimens collected just before extinction, suggesting that the loss of these species was likely extremely rapid (128). Encouragingly, this study also determined that an extant island population long classified as the distinct species *Pseudomys fieldi* is actually a relictual population of the extinct Gould's mouse, *Pseudomys gouldii*, thus taxonomically rescuing *P. gouldii* from extinction.

4.2. Museums as Time Machines: Enabling Robust Investigations of the Anthropocene

In the same way that climate records are used as ground truth in global climate models, the primary evidence needed to validate predictions of how species will respond to future environmental change must come from comparing the historical and current states of a species in relation to environmental change. In this sense, museums can be viewed as eco-evolutionary time machines. Much has been written about the potential for museum collections to act as authoritative records of past biotic assemblages to enable such comparisons. Indeed, there is an increasing number of studies exploiting natural history collections to document and analyze changes in species distributions, phenotype, phenology, and plant–herbivore interactions (71, 98, 135). However, there have been few studies addressing changes in population processes and genetic diversity over time. One exception is the Grinnell Resurvey Project, which exploited the detailed collections of California vertebrates established in the early twentieth century to assess change over a 100-year period in multiple dimensions, including genomic diversity (**Figure 3**), with climate change as a major driver (see the sidebar titled The Grinnell Resurvey Project). In Grinnell's (61, p. 166) prescient words,

> At this point I wish to emphasize what I believe will ultimately prove to be the greatest value of our museum. This value will not, however, be realised until the lapse of many years, possibly a century, assuming that our material is safely preserved. And this is that the student of the future will have access to the original record of faunal conditions in California and the west, wherever we now work.

Looking forward, we will need our eco-evolutionary time machine to enable high-resolution analyses of rapid phenotypic evolution and its constituent genomic contributions. Many analyses and predictions of the capacity for future change rest on a space-for-time substitution, sampling genome diversity across contemporary environmental gradients to evaluate the potential for

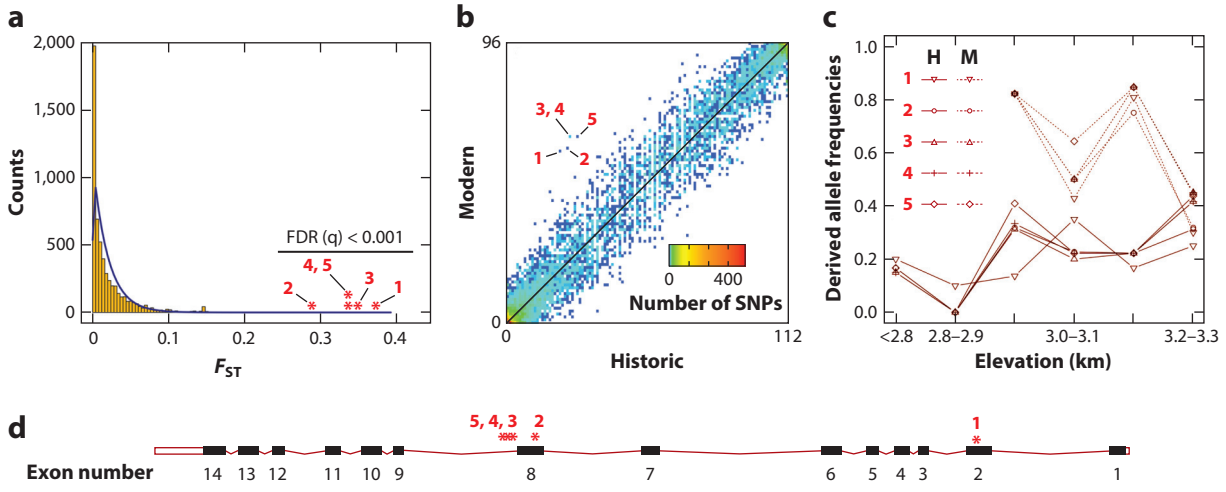


Figure 3

Derived alleles showing significant frequency shifts between historic and modern populations of the alpine chipmunk, *Tamias alpinus*, in Yosemite National Park, California. (*a*) Five outlier SNPs [1–5, FDR (q) < 0.001] are labeled on a plot of the neutral per site temporal F_{ST} distribution (modern versus historic). The histogram of observed F_{ST} (*yellow bins*) is shown with the inferred neutral distribution (*blue line*). (*b*) Unfolded 2D-SFS for SNPs between historic (*x* axis) and modern (*y* axis) Yosemite National Park *Tamias alpinus* specimens. The color of each data point represents the number of SNPs (depicted by the color key) belonging to that particular 2D-SFS category. Leaders point to the five outliers (1–5) showing the only significant allele frequency shifts over time. (*c*) Derived allele frequencies of the five outlier SNPs plotted against sample elevation. Individual sample localities were pooled into 100 m elevational bands to enable allele frequency estimation. (*d*) The position of the five outliers mapped onto the *Alox15* gene in the *Mus musculus* reference genome. SNP 1 is a synonymous mutation (A/G) in exon 2 (chromosome 11:70350801); SNP 2 is synonymous (T/C) and maps to exon 8 (11:70347260); SNPs 3 (T/A), 4 (T/G), and 5 (T/G) are located in the intron between exons 8 and 9. Abbreviations: 2D-SFS, two-dimensional site frequency spectrum; FDR, false discovery rate; F_{ST}, fixation index; SNP, single-nucleotide polymorphism. Figure adapted from Bi et al. (13) (CC BY).

in situ response over time (e.g., 10). Clearly, such approaches using contemporary landscape genomics would benefit if predictions from spatial analyses could be validated from observed changes over time. The ideal (re)sampling approach would entail historical and contemporary collections with dense sampling across space (1, 89) in at least two time points to enable

THE GRINNELL RESURVEY PROJECT

Following the direction given by Grinnell (61), the University of California, Berkeley, Museum of Vertebrate Zoology embarked on a series of centennial resurveys of birds and mammals across elevational transects, mostly in national parks and Forest Service lands in the mountains of California and Nevada. These revealed quite idiosyncratic responses to twentieth-century environmental change (127, 147) even among closely related species. The historical and new collections of specimens of chipmunks (*Tamias*) enabled a multidimensional approach to investigating the ecological and evolutionary processes at play. A species with strong contraction in elevational range (*Tamias alpinus*) underwent selection-driven changes in ecologically relevant skull dimensions (5) and, at microsatellite loci, had reduced allelic diversity and increased genetic structuring of populations as the range became more fragmented (131). Comparative exome sequencing (see **Figure 3**) inferred decreased gene flow among populations and demonstrated rapid evolution in one physiologically relevant gene (13). This evidence and recent resurvey results for desert-adapted species (121) point to the need for integration of genome screens with ecophysiological studies to improve our understanding of how and why species vary in vulnerability to rapid climate change.

detection of changes at landscape scale in phenotype and genotype and in the context of well-documented ecological change.

This represents a challenge to current museum practice, insofar as it will require a combination of (*a*) more intensive collecting at specific locations to enable the diverse requirements of materials for modern genomics and the sample sizes to enable robust estimates of change in phenotype and genotype over time; (*b*) more spatially extensive collecting of tissues to enable robust assessment of how genotypes vary with environment over space (9) and time; and (*c*) improved documentation of sampling methods and ecological conditions, for example, in the form of detailed and accessible field notes, ideally in electronic form (78). All of this points to the need for increased collaboration between field-based ecologists and museum-based curators and consideration of how museums can accession samples and metadata from ecological surveys now and into the future. After all, today's (re)surveys provide the benchmark for future scientists to further understand biotic responses to ongoing change.

DNA methylation: the addition of a methyl group to cytosine or, less commonly, adenine nucleotides; methylation alters gene expression and varies between taxa

Type specimen: the permanent physical specimen on which the description and name of a new species is based

5. MUSEOMICS IN DEEP TIME: ANCIENT DNA

5.1. Development of Ancient DNA

Museum collections have been foundational to the research field known as ancient DNA, which is broadly characterized as genomic research that includes data recovered from historic, degraded samples (140). This field grew out of early experiments sequencing DNA from museum collections, such as a 140-year-old specimen of dried skin from an extinct quagga that was part of the collection at the Natural History Museum Mainz in Germany (69). Since then, and in particular after the advent of high-throughput sequencing technologies that made it possible to amplify fragments of DNA that are too short to be targeted via PCR, most ancient DNA studies either focus entirely on museum-preserved paleontological or archaeological specimens or incorporate specimens that are newly collected from the field that subsequently become part of museum collections.

To date, ancient DNA has been recovered from bones, skin, muscle, teeth, hair, feathers, insect legs, and dental calculus of museum specimens. These investigations have tested hypotheses about systematic relationships of extinct taxa (21, 99, 141), temporal population dynamics of plants and animals (88, 139), human population dispersals (120), dietary preferences (154), and genomic adaptations associated with processes such as domestication (12, 50, 136). Molecular phenotypes have also been assessed, including exploring differential DNA methylation between modern humans and archaic hominins that is potentially associated with diseases and facial phenotypes (58, 59) and assessing temporal changes in DNA methylation in deer mice specimens as old as 76 years (130). The oldest paleontological samples from which DNA has been recovered are an approximately 780,000-year-old horse that was preserved in permafrost in Yukon, Canada (109) and two mammoth bones recovered from permafrost deposits in Siberia that are both older than one million years (150).

5.2. Ancient DNA Can Bolster Inferences from Physical Specimens and Modern Samples

Ancient DNA studies benefit from more than just the preserved DNA in museum samples. Specimens that are part of museum collections are often associated with standardized, accessible metadata that add scientific value to their analysis. For example, the paleontological collection at the American Museum of Natural History in New York contains type specimens for many of the bison species and subspecies named by paleontologists during the eighteenth and early nineteenth

centuries. Comparative analysis of mitochondrial DNA from these specimens and field-collected bison fossils has been used to test hypotheses about the evolutionary relationships among named lineages and to show that many of these are invalid (139, 156). Similarly, ancient DNA research has shown that some descriptions of moa species from New Zealand do not necessarily correspond to their genetic lineages and has provided information on their biology, including demonstrating differential mortality of sexes as the animals grow (2).

Ancient DNA research projects for which the primary materials are field-collected specimens also benefit from comparative data from museums. For instance, genomic analyses of mostly nondiagnostic postcranial remains of horses in Klondike, Yukon, Canada, revealed at least two horse-like lineages present in the region during the last ice age. One lineage was clearly identifiable as caballine horses, *Equus caballus*, but the other had no close genetic match among known taxa. To confirm the hypothesis that this second lineage was the North American stilt-legged horse, *Haringtonhippus francisci*, researchers compared the data to genomic data isolated from a skull in the collection of the Natural History Museum of Los Angeles County that had been identified confidently as a stilt-legged horse. Data from the identifiable skull closely match those from the nondiagnostic postcranial elements, confirming the presence of stilt-legged horses in Yukon (67).

Ancient DNA can also contribute value to museum collections. Some paleontological specimens are too fragmentary to identify with confidence based on morphology or are from individuals at developmental stages that cannot be identified confidently. For instance, ancient DNA revealed that a syntype specimen for the Asian elephant, *Elephas maximus*—a fetus preserved in ethanol that is currently part of the collection of the Swedish Museum of Natural History—is actually an African elephant in the genus *Loxodonta* (25). Similarly, genomic data from a mandible collected from the Taymyr Peninsula, Russia, and accessioned into the paleontology collection of the Zoological Institute of the Russian Academy of Sciences in St. Petersburg as an Arctic wild ass, *Equus hemionus*, indicated that the specimen is in fact a caballine horse, *E. caballus* (151).

Ultraconserved elements (UCEs): short conserved regions of the genome that can be targeted for enrichment and sequencing in large numbers of individuals

5.3. Historical Samples Can Inform Genomic Resource Curation Practices

Finally, ancient DNA can be useful for deriving a more detailed understanding of patterns and rates of DNA degradation over time as samples are stored in natural history collections. Several studies, for example, have sequenced ultraconserved elements (UCEs) from present-day and historical specimens collected up to 150 years ago from taxa including birds (94), bees (14), and daddy longlegs (40). These data collectively show similar slopes of locus fallout over time, although taxon-specific patterns of locus retention are also found (**Figure 4**). Interestingly, one study that sequenced UCEs from 51 carpenter bee (*Xylocopa*) specimens observed that the rate of sample degradation (assessed via DNA concentration and UCE contig length) slowed significantly beginning 21–39 years after preservation (14). Ancient DNA investigations have therefore yielded insights into patterns of DNA degradation over time, which have critical importance for future genomics investigations and museum curatorial practices.

6. CHALLENGES AND OPPORTUNITIES AT THE INTERFACE OF MUSEUMS AND GENOMICS

Museums and the genomics collections they curate have supported a wealth of innovative basic and applied research, yet they face many challenges, and their full potential for supporting research and education has not yet been realized (105). The pressures and insufficiencies facing collections are varied, ranging from gaps in collections coverage to insufficient support, societal misunderstandings of their role, lack of long-term business plans, and challenges with sustaining a diverse and talented pool of staff and curators. Basic training in museum curatorial practices and

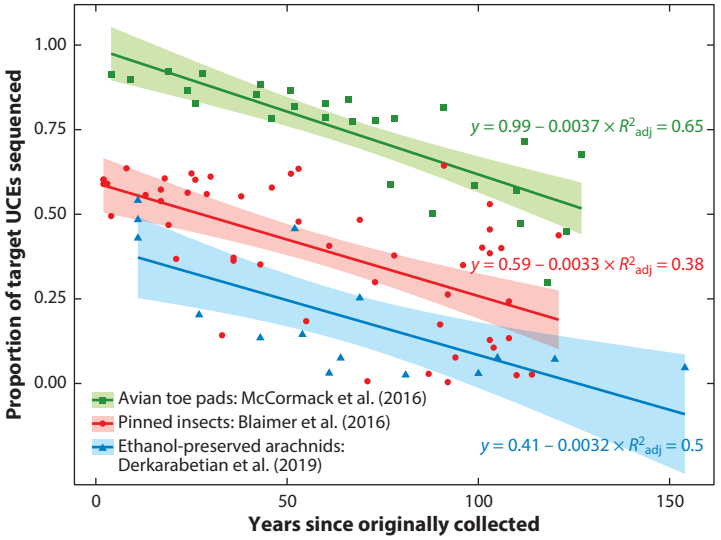


Figure 4

UCE capture success for historical natural history museum specimens in three studies. Comparisons include a study by McCormack et al. (94) using avian toe pad tissue and the Tetrapod 5K probe set (targeting 5,060 UCEs); a study by Blaimer et al. (14) using dry, pinned insects and the Hymenoptera 2.5K probe set (targeting 1,510 UCEs); and a study by Derkarabetian et al. (40) using ethanol-preserved (recent sample in 96% and older samples in 70%) daddy longlegs using the Arachnida 1.1K probe set (targeting 1,120 UCEs). Numbers of loci recovered are scaled based on the total number of UCEs targeted by the specific probe set used. The average percentage of targeted UCE loci that were successfully sequenced decreases by 0.32%, 0.33%, and 0.37% per year for ethanol-preserved arachnids, pinned insects, and dried bird tissue, respectively. Abbreviation: UCE, ultraconserved element. Figure adapted with permission from Derkarabetian et al. (40).

taxonomy can be difficult to find and is in strong decline, especially in some taxonomic areas such as parasitology. There is still no comprehensive compilation of museum genomic collections in the United States or worldwide, although efforts such as iDigBio are helping collate this information (57). The practice of genetic resource and specimen accessioning into museums has waned, resulting in inadequate representation of specimens from the most recent decades in natural history collections, which corresponds to when anthropogenic change is most pronounced (91). A recent National Academies report has highlighted the many pressures facing museum collections today (105). Here, we highlight a few such pressures that are germane to genomics collections specifically.

iDigBio: a National Science Foundation–funded program, based at the University of Florida, to accelerate the digitization of museum specimens and metadata for research and education

Specimen accessioning: the sometimes labor-intensive process by which specimens are added to biological collections, which results in specimens receiving an accession or catalog number

6.1. Data Digitization and Integration

One of the biggest challenges facing the museum community is integrating, maintaining, and sharing the wealth of data associated with natural history collections. This process is far from complete, which constrains museum science investigations, including in areas of genomics, where investigators often lack knowledge about what resources are available for research at natural history museums. Digitization is the process of putting in digital form the metadata associated with museum specimens and, in some cases, digital images of the specimens themselves (65). It is estimated that about 30% of the estimated 800 million to 1 billion US natural history samples have been digitized and made available for searching in electronic databases (105). This shortcoming leaves precious resources inaccessible to researchers and potentially puts these collections

in jeopardy. However, there are also promising models where natural history collections have digitized significant proportions of their collections and made these data widely available online, such as the combined catalog of Australian natural history museums available through the Online Zoological Collections of Australian Museums (OZCAM) (153) and some taxon-specific databases such as VertNet (34). Flexible and extensible database environments, such as Arctos (**https://arctosdb.org/about/details/ecosystem/**), allow for detailed recording of the status of various tissues associated with museum specimens. Still, a major priority for museums is to expand digitization, link diverse collections together into an integrated whole, and make the associated data widely available to all potential stakeholders (6).

VertNet: a National Science Foundation–funded effort to make biodiversity data associated with vertebrates available through digitized, online databases

A related challenge for ongoing digitization is integrating the massive amount of collections data spread across museums and other relevant data providers. Fortunately, global initiatives and their resulting data streams, standards, and tools have made good progress toward this goal. The Global Biodiversity Information Facility (122) is one well-designed online portal for gathering data across natural history collections and enabling global biodiversity syntheses (64, 83). Moreover, several database initiatives have emphasized genomics specifically, including the Global Genome Biodiversity Network (43) and the Global Genome Initiative established by the National Museum of Natural History at the Smithsonian, which collaboratively endeavor to collect, preserve, and disseminate high-quality genomic resources from Earth's biodiversity. Several other broader initiatives have focused on acquiring genomes or transcriptomes of a multitude of organisms in clade-specific investigations in collaboration with major museum providers, such as the Bird 10,000 Genomes (B10K) project, which is actively collaborating with five major museums and several others that provide the raw material for genome sequencing of birds (17).

Other important sources of data typically accumulated by organizations with distinct goals can also be integrated with genomics data from natural history collections. The US National Center for Biotechnology Information (NCBI) and the European Bioinformatics Institute both store massive amounts of molecular data, some of which are directly derived from natural history collections. However, the number of GenBank records that are hyperlinked directly to the record of the museum specimen from which their DNA sequences were derived is still very small. Some database infrastructures, such as Arctos, provide easy generation of URLs for linking genomic data directly with the specimen from which the data were derived (for an example, see **https://mczbase.mcz.harvard.edu/guid/MCZ:Cryo:6597**). The increased presentation of data in supplementary materials of published papers has also had the unfortunate effect of locking away important metadata that could enhance the utility and traceability of genomic data derived from natural history collections. Catalog numbers of specimens used in research papers are often reported, enforced by editors and publishers of journals, particularly in papers stemming from the museum community. But such reporting is often neglected in studies of ancient human remains from museums (e.g., 59). As digitization proceeds, museums, as well as the research and publishing communities, must move to synergize their important collections data with other relevant data sets both for increased scientific impact and traceability and to help ensure long-term persistence.

6.2. Trade-Offs in Genomics and Voucher Specimen Utility

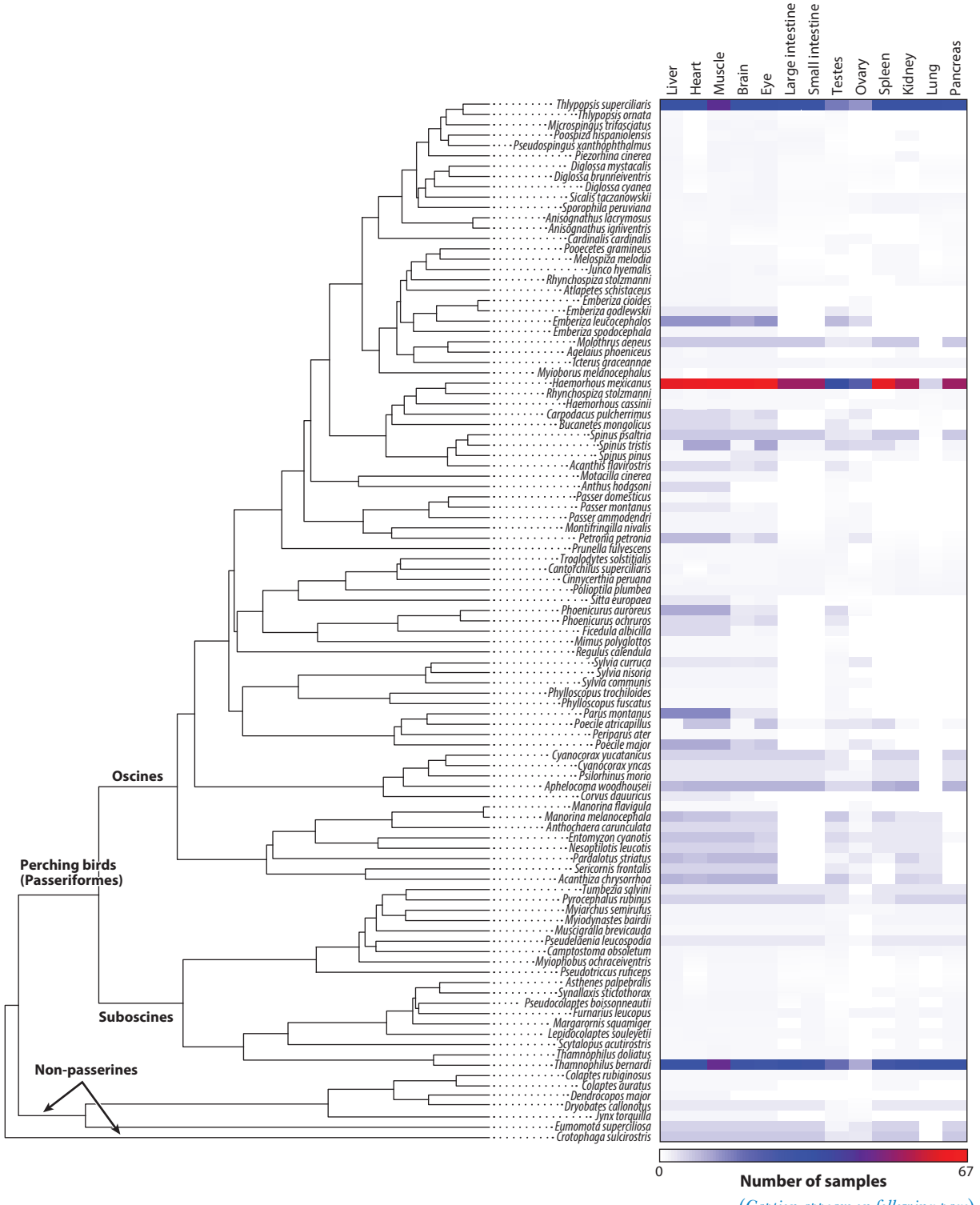
As natural history collections data are digitized and integrated, museums must also work to engage diverse stakeholders who can leverage these resources to make discoveries about our natural world—a process that will likely proceed organically as diverse groups of scientists begin to explore the data available from integrated collections databases. Indeed, genetics and genomics, which were historically more disconnected from museum sciences, provide a powerful example of

how museums can work with interested stakeholders to build new research capacity that unites aspects of both fields, with significant benefits (134). However, key challenges in museum research include how to extract maximum information from finite biological resources and how to serve increasingly diverse research groups addressing broad questions in biology. Collection preservation has numerous potential trade-offs due to the mandate of museums to document biodiversity in all its forms in perpetuity, necessitating diverse curatorial practices, often driven by requirements of individual research projects. For example, some collecting and curatorial practices, such as formalin fixation, are optimized for later extraction of morphological data but function poorly in preserving molecular information and vice versa (95, 129). Diversifying molecular resource collection practices in particular—in terms of tissue types and the kind of genomic assays that could be applied downstream—is an important goal for the museum community, but it can be logistically, financially, and scientifically difficult to implement practices that facilitate a variety of research goals at once.

In practice, general collecting expeditions have prioritized collecting voucher specimens from diverse species, and tissue samples collected for anything but the most standard genomics analyses, such as DNA sequencing, are often relegated to more specialized and targeted fieldwork. Additionally, maximizing tissue quality and type can compromise the quality of the eventual voucher specimen; for example, depending on how one prepares the specimen, extracting RNA-quality brain tissue from a vertebrate specimen can compromise the quality of the resulting skin or skeleton specimen (146), making phenotypic analyses more challenging. Making matters worse, a recent study found that only 11% of 1,300 representative vertebrate genomes available through the NCBI have a clearly referenced voucher specimen (19). Although this rate of vouchering is likely an underestimate, at least for some clades (17), clearly the situation could lead to many potential issues with species identity, sample provenance, verification, and replication. However, targeted expeditions are becoming more common, where the quality and breadth of downstream molecular data are favored over amassing large numbers of specimens and maximizing taxonomic diversity (**Figure 5**). Natural history collections will need to strike a balance between these two collecting paradigms to remain relevant to the broadest array of researchers wishing to use genomic resource collections.

6.3. Updating Cryopreservation Standards

Looking forward, new collaborations built on a foundation of genomics will emerge between museum scientists and investigators from other fields like medicine, engineering, chemistry, and technology. Arguably, engaging diverse stakeholders sooner rather than later will be important for adapting collection and curatorial procedures for broader use (6, 60, 105, 134, 142). Unfortunately, large proportions of tissue collections and most voucher specimens in many museums were collected in an era that demanded minimum quality standards in terms of macromolecule integrity. The low quality of many macromolecules in museum cryogenic collections has meant that studies aiming to embrace long-read, transcriptome, or epigenome sequencing, for instance, may need to resort to alternative sources of genomic material for immediate use. This need for new collections helps to explain why many genomics studies rely instead on freshly-collected materials; however, these new collections are often not deposited in natural history collections for long-term preservation. Even for fossil specimens, there is a suite of best practices to ensure the longevity of macromolecules, but implementing these sometimes stringent requirements can require infrastructure traditionally beyond the purview of many natural history collections. Some paleontological collections have begun storing specimens in environmentally controlled rooms, at high expense, such as the collections of the Museum of the North in Fairbanks, Alaska, and the Canadian Museum of Nature in Ottawa.



(*Caption appears on following page*)

Figure 5 (*Figure appears on preceding page*)

An example of next-generation, RNA-ready genomic samples for diverse avian species in the Museum of Comparative Zoology (MCZ) at Harvard University. The samples shown have been collected since 2012 from diverse locations around the globe and are compatible with RNA-seq and other methods utilizing flash-frozen tissues. Most samples were minced and preserved in RNAlater within 10 min of sacrifice and then flash frozen in liquid nitrogen after ~12 h at cool temperatures. A small percentage of samples were directly flash frozen without RNAlater. All samples are now permanently stored in liquid nitrogen and queryable through the MCZbase database, and a subset has been tested to verify the presence of high-quality RNA for gene expression or genome annotation purposes. Gustavo Bravo and Jonathan Schmitt retrieved and analyzed data and made an early draft of this figure. All counts of next-generation, RNA-ready genomics samples are available online at **https://doi.org/10.5281/zenodo.5093840** and **https://edwards-bird-lab.github.io/museum-genomics/**.

Unfortunately, rigorous investigations of how to optimally preserve, store, extract, and utilize macromolecules, such as DNA and RNA, from tissue samples are rare, leaving museum practitioners with only anecdotal information about ideal preservation conditions for molecular samples. This shortcoming represents a major gap in understanding that must be addressed and disseminated across collections so that new, improved curatorial practices can be implemented and standardized for the benefit of the broader research community. Dialog with related collections institutions, like biobanks used in human medical research, would benefit the museum genomics community (92). Once in place, these practices can be adopted when collecting new genomic samples, but we emphasize that collections should also target existing holdings for curatorial improvements. As a case in point, the Museum of Comparative Zoology at Harvard has received funds from the US National Science Foundation (NSF) to subsample type specimens from all its collections and store them in the museum's cryogenic collection, with the hope of being able to sequence the genomes of all these types in the future. Initiatives such as the Frozen Zoo at the San Diego Zoo Wildlife Alliance (**https://science.sandiegozoo.org/resources/frozen-zoo®**) have been archiving viable cell cultures, which can be used to repeatedly grow cell stocks and enable improved preservation of cellular biomolecules, functional studies, and even the possibility of eventually contributing to the conservation of rare or extinct species (107, 132, 157). Increasingly, curators are requiring the sequencing of whole genomes from rare or fragile specimens so as to complete the practical genomic inventory of those specimens and minimize or eliminate additional subsampling in the future.

6.4. Credit and Data Attribution

The finite nature of natural history collections and the large amounts of time, effort, and funding that go into building and preserving collections can also result in unaligned, or even conflicting, research priorities. In particular, whereas genetics and genomics have revolutionized biodiversity sciences in many ways, there continues to be a gap, although narrowing, between biodiversity approaches focusing on phenotypes and those focusing on genotypes. Conflicts can therefore arise over how materials are collected and preserved, which uses are appropriate and have priority, and how investigators should share credit for scientific discoveries. Unfortunately, the protocols for attribution of data and assignment of credit for specimen acquisition in biodiversity studies are still very rudimentary, leading to loss of data traceability through the cycle from field to museum to publication. As a result, museum staff and field collectors often receive little credit, implicitly or explicitly, for studies analyzing museum specimens collected by them or in their care (6, 126). Many museum staff are neutral about being credited with a data source in their care, whereas others legitimately require coauthorship on studies analyzing or using specimens in their care.

Data digitization and integration will help track specimen use in detail and allow better attribution and documentation of credit to the numerous field biologists and curators who collected, prepared, identified, and accessioned specimens being used in genomics research (134, 162).

Moreover, international agreements can also function to establish best practices for sharing museum resources and data, especially in the international context. The Nagoya Protocol (138; see also **https://www.cbd.int/abs/**), for example, is setting a new standard for access to museum specimens and, potentially, digital data stemming from genomic analysis. Overall, the museum community generally acknowledges the continued need for such protocols after decades of sometimes less-than-satisfactory inclusion or acknowledgment of researchers, infrastructure, and investment (33). Museum collections are increasingly communal, and although exclusive access to certain materials may be appropriate over reasonable timescales—an arrangement that hinges critically on how the collection and research are funded—all resources should ultimately be available to the entire research community.

Nagoya Protocol: an international agreement that aims to ensure fair and equitable utilization of genetic resources for the benefit of mankind

Natural Science Collections Alliance: a nonprofit association that supports natural science collections, enabling diverse research activities for the benefit of science and society

6.5. Cost Recovery of Genomics Collections

Different collections take alternative approaches to address the persistent challenge of recovering costs for serving the genomics community, as well as ensuring their own long-term livelihood, although some attempts at articulating universal best practices are appearing (126). Some museum genomics collections, citing the significant costs of collecting, maintaining, and sharing physical museum resources, charge outside researchers fees for assembling and sending off genomics samples for external use. As pressing as cost recovery is, the need for museums to instill a culture of accountability and reciprocity among potential users of collections is important. For instance, as part of a typical request to use museum genomic resources, many museums require attestation of how the requester has contributed to the broader museum mission, either through their own field collecting or specimen preparation, to engender a greater appreciation of the considerable time, cost, and effort required to amass such collections. Researchers requesting materials from museums for destructive sampling typically are asked to provide detailed descriptions of experimental procedures so that museum staff can ascertain whether best practices will be adopted before precious genomics samples are destroyed. Such policies are applied haphazardly across institutions and research programs, and standardization of best practices for requesting, obtaining, and accessioning museum resources will help to overcome these difficulties (105). Funding bodies requiring specimen management plans, analogous to existing data management plans for the NSF and National Institutes of Health, could help speed adoption of best practices and ensure that the often considerable costs of accessioning and curation are adequately covered by research programs hoping to archive specimens in a museum or make use of existing collections (35).

Finally, the economic vibrancy and fragility of museums and their collections vary greatly across countries, institutional context, and the public-private continuum, and this variation complicates the ways in which museums can help advance and integrate genomic science (105). Several societies and consortia, such as the Natural Science Collections Alliance and the American Institute of Biological Sciences, provide advocacy and support for museum collections, particularly in the United States. Major historical losses, such as the physical damage sustained by the Museum für Naturkunde in Berlin during World War II, as well as recent catastrophes, such as the fires gutting several major museums and collections in Brazil (125) and narrowly missing a major botanical collection in South Africa (106), underscore the fragile security and infrastructure behind many museums and the need for duplication and offsite replication of specimens and data. Many herbaria routinely split parts of a single organism and distribute it across multiple institutions for preservation and care, a practice that is increasingly undertaken by genomics collections, especially in the context of international collaborations. Such a tradition greatly improves the chances of parts of that individual surviving a catastrophic collapse at one institution and also makes the individual specimen more accessible to researchers around the world. Such practices are particularly

important, and increasingly required, for archiving type specimens, the original specimen on which a species description is based. Such data and specimen redundancy ensure the longevity of data and genomic resources linked to individual specimens—a crucial backup in a world in which even national museums are poorly funded and understaffed (6, 105).

6.6. Ethics of Museum Genomics

Museums and museum collections have recently begun to reckon with a past that in many cases is closely intertwined with a history of colonial exploitation and racism (39). The challenge of distinguishing scientific acts of specimen collection from an often unsavory history of colonial dominance is arguably more difficult for anthropological collections, but genomics collections are also not immune, especially in the realm of ancient DNA of humans. For example, the first genome of an Indigenous Australian was sequenced from a lock of hair accessioned in the collection of the Leverhulme Centre for Human Evolutionary Studies in Cambridge, United Kingdom (118). Although cleared by ethical review boards in Denmark, where most of the research was undertaken, the study nonetheless raised concerns among scientists and was acknowledged to be in "uncharted ethical territory" (23, p. 522). Our focus in this review is not on human evolutionary studies, but zoological specimens and cryogenic collections in zoological museums also reveal ethical challenges of acquisition and use. In some countries, such as New Zealand, study, collection, transport, or manipulation of indigenous fauna requires the consent of representatives of indigenous peoples. In many countries, conducting zoological research on indigenous lands also requires multiple permissions and consultations. The Nagoya Protocol is written and being refined in part to include indigenous peoples in the process of prioritizing research and gaining access to its intellectual and economic benefits.

Although there are few examples thus far of ethical dilemmas specifically at the intersection of museums and cryogenic collections, the museum genomics community is beginning to articulate more precise ethical and cultural standards, taking cues from more developed protocols from human genomics (29). Field expeditions from many museums have traditionally included, or are increasingly required to include, in-country collaborators, not merely as assistants but as coauthors and beneficiaries of advanced training and knowledge transfer. Recent genome projects focusing on culturally sensitive species have included indigenous consultation from project inception and shared authorship (54). The potential drawbacks of such restrictions for both in- and out-country scientists cannot be underestimated. Regulations in some countries, such as India, which severely restrict or do not permit conducting genetic studies in foreign laboratories, or require that DNA be destroyed and not archived after a study has been completed, decrease opportunities for responsible archiving, international collaborations, and the technology transfer and training that often ensues. In addition, some nations have strict legislation about returning type material that is described abroad to their museums, something that museums need to take into consideration in the long term because the average time between specimen collection and its scientific description is about 21 years (49). The ethics of museum genomics, and of the museum enterprise in general, is ripe for active discussion and updating, a process that will hopefully itself be inclusive and widely consultative.

7. MUSEUMS AND GENOMICS IN THE FUTURE

The archiving of genomic resources and the effort to make these accessible to the scientific community represent some of the most recent paradigm shifts in our ability to extract information from museum specimens. Future technological advancements will undoubtedly shape museum sciences in diverse, unpredictable ways. As such, the preservation of existing natural

history collections, which remains a major challenge at most institutions, is of vital importance for ensuring that researchers can leverage scientific collections to answer a range of important biological questions. Of equal importance will be the scientifically driven, sustainable growth of these collections in the future, a growth that will advance new areas of collections-based research to benefit science, education, conservation, and humanity (91).

Ambitious efforts to generate high-quality genome assemblies for many—if not all—species are now underway with projects like the Vertebrate Genomes Project (**https://vertebrategenomesproject.org/**), B10K, the Earth BioGenome Project (86), components of European Commission programs like Horizon Europe (**https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en**), and the Wellcome Trust–funded Tree of Life project headed by the Sanger Institute in the United Kingdom (**https://www.sanger.ac.uk/programme/tree-of-life/**). To varying degrees, these projects are working with museums to guide taxon sampling and sequence as much as possible from archived voucher specimens. Other projects, like DNAmark in Denmark or the National Ecological Observatory Network in the United States, are focusing primarily on eDNA and are causing museums to consider how to archive environmental samples in more of an Earth-monitoring vein (35). The emergence of new sequence-based assays, particularly those tailored to old, frozen, or degraded samples (32, 36), will unlock access to new pieces of genomic and cellular information. Already, a range of functional genomics assays is in place that enables profiling of epigenetic states (8, 85), with new advancements being regularly made (e.g., 145, 158). Unfortunately, aside from the studies of methylation in Neanderthals and deer mice discussed in Section 5.1 (58, 130), functional genomics approaches have not yet begun leveraging museum genomics collections. Analogous fields to genomics that sequence proteins or metabolic products —known as proteomics and metabolomics, respectively—have already begun to benefit from museum genomics collections (4, 108, 137, 155), and with more maturation, these methods will eventually provide the potential for data extraction that is on par with current genomics methodologies.

Finally, while further afield from traditional museum science, we predict that the genome editing revolution, driven primarily by the discovery and development of CRISPR-Cas9 editing systems, will power new research programs based in natural history museums (16). Indeed, as genetic editing is applied more widely, including within nontraditional model species (37, 119), museums may benefit greatly by positioning themselves as important repositories of genetically modified organisms, enabling integration with other fields of biology and science that have been relatively estranged from traditional museum studies. Altogether, museum environments may become important catalysts for increased cohesiveness between genomics and other ongoing technological advancements that will benefit museum scientists, including imaging [e.g., whole-body computed tomography (CT) scanning or cellular and nuclear imaging via advanced microscopy (51, 55)], computer science [e.g., big data analysis or artificial intelligence (161)], and engineering [e.g., robotics and nature-inspired designs (60)]. However, despite all of these possible research directions, natural history museums must also continue to champion the importance of collections and basic natural history research to newly engaged scientists, government agencies, nonprofit groups, and the public through well-designed outreach efforts.

In the future, connections between museum science and genomics will grow stronger through the adoption of curatorial practices facilitating the use of cutting-edge techniques. For example, single-cell and functional genomics approaches offer one particularly fruitful area of research capable of providing exponential increases in the resolution of information that can be extracted from museum specimens (62, 93), and there may be manageable protocols allowing such methods to be applied to archived tissues. Innovations in genomics technologies will expand the kinds of

genomic data that can be gleaned from museum collections, but many of these new data streams will fall outside traditional areas of research by museum curators as scientists clamor to tap the information reserves stored in natural history collections. The expertise of new, diverse stakeholders, including researchers in medicine, engineering, and other fields, will need to be engaged and consulted to modernize curatorial practices in museum genomics and to increase the utility of genomics collections, thereby solidifying and expanding the important role that museums play in science. Increased digitization and integration of global museum collections will be crucial to this endeavor. Future collection, curatorial, and data collection approaches, including those leveraging genomics, will unlock critical pieces of ecological, evolutionary, or conservation information, but only insofar as museum scientists succeed in gathering baseline measures of biodiversity worldwide (91). Overall, while these research themes represent some logical directions for museum-based investigations over the next half century, numerous equally important and unique research projects will also emerge that will be capable of revolutionizing museum science and illuminating diverse biological principles.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank the curators and collections managers at several institutions for assistance and discussions in quantifying temporal trends in cryogenic collection growth: Cheryl Hayashi (American Museum of Natural History); Christopher Austin, Robb Brumfield, Prosanta Chakrabarty, and Donna Dittmann (Louisiana Museum of Natural History); Breda Zimkus and Brendan Haley (Museum of Comparative Zoology); Michael Anderson, Joseph Cook, Thomas Turner, and Christopher Witt (Museum of Southwestern Biology); and Carla Cicero, Michelle Koo, and Carol Spencer (Museum of Vertebrate Zoology). Shahan Derkarabetian kindly provided the raw data for **Figure 4**. Gustavo Bravo and Jonathan Schmitt provided the raw data and initial draft for **Figure 5** and helpful discussions of optimal collecting approaches for museum genomics. We thank Carla Cicero, Ian Owens, Breda Zimkus, and Hank Bart for helpful discussion of issues related to this manuscript, and John Bates for helpful comments on the manuscript. Work in the lab of S.V.E. is supported by the National Science Foundation (NSF), most recently grant DEB-1831560. D.C.C is supported by a postdoctoral fellowship from the NSF (DEB-1812310). This work was supported by the NSF "Natural History: Preserving the Genomes of the Type Specimens in the Museum of Comparative Zoology" grant to G.G. and S.V.E. (CSBR-1946857) and by an award from the Institute for Museum and Library Science to B.S. (MG-30-17-0045-17). Research in the lab of C.M. is supported by the Australian Research Council.

LITERATURE CITED

1. Ahrens CW, Rymer PD, Stow A, Bragg J, Dillon S, et al. 2018. The search for loci under selection: trends, biases and progress. *Mol. Ecol.* 27(6):1342–56
2. Allentoft ME, Bunce M, Scofield RP, Hale ML, Holdaway RN. 2010. Highly skewed sex ratios and biased fossil deposition of moa: Ancient DNA provides new insight on New Zealand's extinct megafauna. *Quat. Sci. Rev.* 29(5):753–62
3. Alves JM, Carneiro M, Cheng JY, de Matos AL, Rahman MM, et al. 2019. Parallel adaptation of rabbit populations to myxoma virus. *Science* 363(6433):1319–26

4. Asara JM, Schweitzer MH, Freimark LM, Phillips M, Cantley LC. 2007. Protein sequences from mastodon and *Tyrannosaurus rex* revealed by mass spectrometry. *Science* 316(5822):280–85
5. Assis APA, Patton JL, Hubbe A, Marroig G. 2016. Directional selection effects on patterns of phenotypic (co)variation in wild populations. *Proc. R. Soc. B* 283(1843):20161615
6. Bakker FT, Antonelli A, Clarke JA, Cook JA, Edwards SV, et al. 2020. The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ* 8:e8225
7. Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, et al. 2011. Has the Earth's sixth mass extinction already arrived? *Nature* 471(7336):51–57
8. Barrera-Redondo J, Piñero D, Eguiarte LE. 2020. Genomic, transcriptomic and epigenomic tools to study the domestication of plants and animals: a field guide for beginners. *Front. Genet.* 11:742
9. Battey CJ, Ralph PL, Kern AD. 2020. Space is the place: effects of continuous spatial structure on analysis of population genetic data. *Genetics* 215(1):193–214
10. Bay RA, Harrigan RJ, Underwood VL, Gibbs HL, Smith TB, Ruegg K. 2018. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* 359(6371):83–86
11. Bell KC, Carlson CJ, Phillips AJ. 2018. Parasite collections: overlooked resources for integrative research and conservation. *Trends Parasitol.* 34(8):637–39
12. Bergström A, Frantz L, Schmidt R, Ersmark E, Lebrasseur O, et al. 2020. Origins and genetic legacy of prehistoric dogs. *Science* 370(6516):557–64
13. Bi K, Linderoth T, Singhal S, Vanderpool D, Patton JL, et al. 2019. Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLOS Genet.* 15(5):e1008119
14. Blaimer BB, Lloyd MW, Guillory WX, Brady SG. 2016. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLOS ONE* 11(8):e0161531
15. Blom MPK. 2021. Opportunities and challenges for high-quality biodiversity tissue archives in the age of long-read sequencing. *Mol. Ecol.* In press. **https://doi.org/10.1111/mec.15909**
16. Bono JM, Olesnicky EC, Matzkin LM. 2015. Connecting genotypes, phenotypes and fitness: harnessing the power of CRISPR/Cas9 genome editing. *Mol. Ecol.* 24(15):3810–22
17. Bravo GA, Schmitt CJ, Edwards SV. 2021. What have we learned from the first 500 avian genomes? *Annu. Rev. Ecol. Evol. Syst.* 52:611–39
18. Brealey JC, Leitão HG, van der Valk T, Xu W, Bougiouri K, et al. 2020. Dental calculus as a tool to study the evolution of the mammalian oral microbiome. *Mol. Biol. Evol.* 37(10):3003–22
19. Buckner JC, Sanders RC, Faircloth BC, Chakrabarty P. 2021. The critical importance of vouchers in genomics. *eLife* 10:e68264
20. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10(12):1213–18
21. Bunce M, Szulkin M, Lerner HRL, Barnes I, Shapiro B, et al. 2005. Ancient DNA provides new insights into the evolutionary history of New Zealand's extinct giant eagle. *PLOS Biol.* 3(1):e9
22. Burrell AS, Disotell TR, Bergey CM. 2015. The use of museum specimens with high-throughput DNA sequencers. *J. Hum. Evol.* 79:35–44
23. Callaway E. 2011. Aboriginal genome analysis comes to grips with ethics. *Nature* 477(7366):522–23
24. Campos PF, Craig OE, Turner-Walker G, Peacock E, Willerslev E, Gilbert MTP. 2012. DNA in ancient bone—Where is it located and how should we extract it? *Ann. Anat.* 194(1):7–16
25. Cappellini E, Gentry A, Palkopoulou E, Ishida Y, Cram D, et al. 2014. Resolution of the type material of the Asian elephant, *Elephas maximus* Linnaeus, 1758 (Proboscidea, Elephantidae). *Zool. J. Linn. Soc.* 170(1):222–32
26. Carvalho T, Becker CG, Toledo LF. 2017. Historical amphibian declines and extinctions in Brazil linked to chytridiomycosis. *Proc. R. Soc. B* 284(1848):20162254
27. Cheng S, Melkonian M, Smith SA, Brockington S, Archibald JM, et al. 2018. 10KP: A phylodiverse genome sequencing plan. *GigaScience* 7:giy013
28. Cheviron ZA, Carling MD, Brumfield RT. 2011. Effects of postmortem interval and preservation method on RNA isolated from field-preserved avian tissues. *Condor* 113(3):483–89

29. Claw KG, Anderson MZ, Begay RL, Tsosie KS, Fox K, Garrison NA. 2018. A framework for enhancing ethical genomic research with Indigenous communities. *Nat. Commun.* 9(1):2957
30. Colella JP, Agwanda BR, Khan FAA, Bates J, Carrion Bonilla CA, et al. 2020. Build international biorepository capacity. *Science* 370(6518):773–74
31. Colella JP, Bates J, Burneo SF, Camacho MA, Carrion Bonilla C, et al. 2021. Leveraging natural history biorepositories as a global, decentralized, pathogen surveillance network. *PLOS Pathog.* 17(6):e1009583
32. Colella JP, Tigano A, MacManes MD. 2020. A linked-read approach to museomics: higher quality de novo genome assemblies from degraded tissues. *Mol. Ecol. Resour.* 20(4):856–70
33. Colwell C. 2016. Collaborative archaeologies and descendant communities. *Annu. Rev. Anthropol.* 45:113–27
34. Constable H, Guralnick R, Wieczorek J, Spencer C, Peterson AT, VertNet Steer. Comm.. 2010. VertNet: a new model for biodiversity data sharing. *PLOS Biol.* 8(2):e1000309
35. Cook JA, Greiman SE, Agosta SJ, Anderson RP, Arbogast BS, et al. 2016. Transformational principles for NEON sampling of mammalian parasites and pathogens: a response to Springer and colleagues. *BioScience* 66(11):917–19
36. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, et al. 2017. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* 14(10):959–62
37. Crawford K, Quiroz JFD, Koenig KM, Ahuja N, Albertin CB, Rosenthal JJC. 2020. Highly efficient knockout of a squid pigmentation gene. *Curr. Biol.* 30(17):3484–3490.e4
38. Crutzen PJ. 2002. Geology of mankind. *Nature* 415(6867):23
39. Das S, Lowe M. 2018. Nature read in black and white: decolonial approaches to interpreting natural history collections. *J. Nat. Sci. Collect.* 6:4–14
40. Derkarabetian S, Benavides LR, Giribet G. 2019. Sequence capture phylogenomics of historical ethanol-preserved museum specimens: unlocking the rest of the vault. *Mol. Ecol. Resour.* 19(6):1531–44
41. Dessauer HC, Hafner MS. 1984. *Collections of Frozen Tissues: Value, Management, Field and Laboratory Procedures, and Directory of Existing Collections*. Lawrence, KS: Assoc. Syst. Collect.
42. DiEuliis D, Johnson KR, Morse SS, Schindel DE. 2016. Specimen collections should have a much bigger role in infectious disease research and response. *PNAS* 113(1):4–7
43. Droege G, Barker K, Seberg O, Coddington J, Benson E, et al. 2016. The Global Genome Biodiversity Network (GGBN) Data Standard specification. *Database* 2016:baw125
44. Dunnum JL, Yanagihara R, Johnson KM, Armien B, Batsaikhan N, et al. 2017. Biospecimen repositories and integrated databases as critical infrastructure for pathogen discovery and pathobiology research. *PLOS Negl. Trop. Dis.* 11(1):e0005133
45. Edwards SV, Birks S, Brumfield RT, Hanner R. 2005. Future of avian genetic resources collections: archives of evolutionary and environmental history. *Auk* 122(3):979–84
46. Eldridge MDB, Deakin JE, MacDonald AJ, Byrne M, Fitzgerald A, et al. 2020. The Oz Mammals Genomics (OMG) initiative: developing genomic resources for mammal conservation at a continental scale. *Aust. Zool.* 40(3):505–9
47. Field KG, Olsen GJ, Lane DJ, Giovannoni SJ, Ghiselin MT, et al. 1988. Molecular phylogeny of the animal kingdom. *Science* 239(4841):748–53
48. Fong JJ, Cheng TL, Bataille A, Pessier AP, Waldman B, Vredenburg VT. 2015. Early 1900s detection of *Batrachochytrium dendrobatidis* in Korean amphibians. *PLOS ONE* 10(3):e0115656
49. Fontaine B, Perrard A, Bouchet P. 2012. 21 years of shelf life between discovery and description of new species. *Curr. Biol.* 22(22):R943–44
50. Frantz LAF, Bradley DG, Larson G, Orlando L. 2020. Animal domestication in the era of ancient genomics. *Nat. Rev. Genet.* 21(8):449–60
51. Friedrich F, Matsumura Y, Pohl H, Bai M, Hörnschemeyer T, Beutel RG. 2014. Insect morphology in the age of phylogenomics: innovative techniques and its future role in systematics. *Entomol. Sci.* 17(1):1–24
52. Friis G, Fandos G, Zellmer AJ, McCormack JE, Faircloth BC, Milá B. 2018. Genome-wide signals of drift and local adaptation during rapid lineage divergence in a songbird. *Mol. Ecol.* 27(24):5137–53

53. Gauthier J, Pajkovic M, Neuenshwander S, Kaila L, Schmid S, et al. 2020. Museomics identifies genetic erosion in two butterfly species across the 20th century in Finland. *Mol. Ecol. Resour.* 20(5):1191–205
54. Gemmell NJ, Rutherford K, Prost S, Tollis M, Winter D, et al. 2020. The tuatara genome reveals ancient features of amniote evolution. *Nature* 584(7821):403–9
55. Gignac PM, Kley NJ, Clarke JA, Colbert MW, Morhardt AC, et al. 2016. Diffusible iodine-based contrast-enhanced computed tomography (diceCT): an emerging tool for rapid, high-resolution, 3-D imaging of metazoan soft tissues. *J. Anat.* 228(6):889–909
56. Gilbert MTP, Moore W, Melchior L, Worobey M. 2007. DNA extraction from dry museum beetles without conferring external morphological damage. *PLOS ONE* 2(3):e272
57. Godden GT, Soltis P. 2014. A new iDigBio web feature links DNA banks and genetic resources repositories in the United States. In *DNA Banking for the 21st Century: Proceedings of the U.S. Workshop on DNA Banking*, ed. WL Applequist, LM Campbell, pp. 173–81. St. Louis, MO: Missouri Botanical Garden
58. Gokhman D, Lavi E, Prüfer K, Fraga MF, Riancho JA, et al. 2014. Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. *Science* 344(6183):523–27
59. Gokhman D, Nissim-Rafinia M, Agranat-Tamir L, Housman G, García-Pérez R, et al. 2020. Differential DNA methylation of vocal and facial anatomy genes in modern humans. *Nat. Commun.* 11(1):1189
60. Green DW, Watson JA, Jung H-S, Watson GS. 2019. Natural history collections as inspiration for technology. *BioEssays* 41(2):e1700238
61. Grinnell J. 1910. The methods and uses of a research museum. *Pop. Sci. Mon.* 77:163–69
62. Hahn EE, Grealy A, Alexander M, Holleley CE. 2020. Museum epigenomics: charting the future by unlocking the past. *Trends Ecol. Evol.* 35(4):295–300
63. Harvey MG, Bravo GA, Claramunt S, Cuervo AM, Derryberry GE, et al. 2020. The evolution of a tropical biodiversity hotspot. *Science* 370(6522):1343–48
64. Heberling JM, Miller JT, Noesgaard D, Weingart SB, Schigel D. 2021. Data integration enables global biodiversity synthesis. *PNAS* 118(6):e2018093118
65. Hedrick BP, Heberling JM, Meineke EK, Turner KG, Grassa CJ, et al. 2020. Digitization and the future of natural history collections. *BioScience* 70(3):243–51
66. Heindler FM, Christiansen H, Frédérich B, Dettaï A, Lepoint G, et al. 2018. Historical DNA metabarcoding of the prey and microbiome of trematomid fishes using museum samples. *Front. Ecol. Evol.* 6:151
67. Heintzman PD, Zazula GD, MacPhee RD, Scott E, Cahill JA, et al. 2017. A new genus of horse from Pleistocene North America. *eLife* 6:e29944
68. Hickey JJ, Anderson DW. 1968. Chlorinated hydrocarbons and eggshell changes in raptorial and fish-eating birds. *Science* 162(3850):271–73
69. Higuchi R, Bowman B, Freiberger M, Ryder OA, Wilson AC. 1984. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312(5991):282–84
70. Hofreiter M. 2012. Nondestructive DNA extraction from museum specimens. In *Ancient DNA: Methods and Protocols*, ed. B Shapiro, M Hofreiter, pp. 93–100. Totowa, NJ: Humana Press
71. Holmes MW, Hammond TT, Wogan GOU, Walsh RE, LaBarbera K, et al. 2016. Natural history collections as windows on evolutionary processes. *Mol. Ecol.* 25(4):864–81
72. Hope AG, Sandercock BK, Malaney JL. 2018. Collection of scientific specimens: benefits for biodiversity sciences and limited impacts on communities of small mammals. *BioScience* 68(1):35–42
73. Hydeman ME, Longo AV, Velo-Antón G, Rodriguez D, Zamudio KR, Bell RC. 2017. Prevalence and genetic diversity of *Batrachochytrium dendrobatidis* in Central African island and continental amphibian communities. *Ecol. Evol.* 7(19):7729–38
74. i5K Consortium. 2013. The i5K Initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment. *J. Hered.* 104(5):595–600
75. Jackson JA, Laikre L, Baker CS, Kendall KC, Genet. Monit. Work. Group. 2012. Guidelines for collecting and maintaining archives for genetic monitoring. *Conserv. Genet. Resour.* 4(2):527–36
76. Johnson KG, Brooks SJ, Fenberg PB, Glover AG, James KE, et al. 2011. Climate change and biosphere response: unlocking the collections vault. *BioScience* 61(2):147–53
77. Johnson NK, Zink RM, Barrowclough GF, Marten JA. 1984. Suggested techniques for modern avian systematics. *Wilson Bull.* 96(4):543–60

78. Jones M. 2017. From personal to public: field books, museums, and the opening of the archives. *Arch. Rec.* 38(2):212–27
79. Joseph L. 2011. Museum collections in ornithology: today's record of avian biodiversity for tomorrow's world. *Emu Austral. Ornithol.* 111(3):i–xii
80. Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, et al. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *PNAS* 86(16):6196–200
81. Kohn MH, Pelz H-J, Wayne RK. 2000. Natural selection mapping of the warfarin-resistance gene. *PNAS* 97(14):7911–15
82. Krell F-T, Wheeler QD. 2014. Specimen collection: Plan for the future. *Science* 344(6186):815–16
83. La Salle J, Williams KJ, Moritz C. 2016. Biodiversity analysis in the digital era. *Philos. Trans. R. Soc. B* 371:20150337
84. Lamichhaney S, Card DC, Grayson P, Tonini JFR, Bravo GA, et al. 2019. Integrating natural history collections and comparative genomics to study the genetic architecture of convergent evolution. *Philos. Trans. R. Soc. B* 374(1777):20180248
85. Lappalainen T. 2015. Functional genomics bridges the gap between quantitative genetics and molecular biology. *Genome Res.* 25(10):1427–31
86. Lewin HA, Robinson GE, Kress WJ, Baker WJ, Coddington J, et al. 2018. Earth BioGenome Project: sequencing life for the future of life. *PNAS* 115(17):4325–33
87. Lister AM. 2011. Natural history collections as sources of long-term datasets. *Trends Ecol. Evol.* 26(4):153–54
88. Lorenzen ED, Nogués-Bravo D, Orlando L, Weinstock J, Binladen J, et al. 2011. Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479(7373):359–64
89. Lotterhos KE, Whitlock MC. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* 24(5):1031–46
90. Mack KL, Ballinger MA, Phifer-Rixey M, Nachman MW. 2018. Gene regulation underlies environmental adaptation in house mice. *Genome Res.* 28(11):1636–45
91. Malaney JL, Cook JA. 2018. A perfect storm for mammalogy: declining sample availability in a period of rapid environmental degradation. *J. Mammal.* 99(4):773–88
92. Malsagova K, Kopylov A, Stepanov A, Butkova T, Sinitsyna A, et al. 2020. Biobanks—a platform for scientific and biomedical research. *Diagnostics* 10(7):485
93. Marioni JC, Arendt D. 2017. How single-cell genomics is changing evolutionary and developmental biology. *Annu. Rev. Cell Dev. Biol.* 33:537–53
94. McCormack JE, Tsai WLE, Faircloth BC. 2016. Sequence capture of ultraconserved elements from bird museum specimens. *Mol. Ecol. Resour.* 16(5):1189–203
95. McGuire JA, Cotoras DD, O'Connell B, Lawalata SZS, Wang-Claypool CY, et al. 2018. Squeezing water from a stone: high-throughput sequencing from a 145-year old holotype resolves (barely) a cryptic species problem in flying lizards. *PeerJ* 6:e4470
96. McKenzie JA, Batterham P. 1994. The genetic, molecular and phenotypic consequences of selection for insecticide resistance. *Trends Ecol. Evol.* 9(5):166–69
97. Medlin L, Elwood HJ, Stickel S, Sogin ML. 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71(2):491–99
98. Meineke EK, Davies TJ, Daru BH, Davis CC. 2019. Biological collections for understanding biodiversity in the Anthropocene. *Philos. Trans. R. Soc. B* 374(1763):20170386
99. Meyer M, Palkopoulou E, Baleka S, Stiller M, Penkman KEH, et al. 2017. Palaeogenomes of Eurasian straight-tusked elephants challenge the current view of elephant evolution. *eLife* 6:e25413
100. Moore G, Tessler M, Cunningham SW, Betancourt J, Harbert R. 2020. Paleo-metagenomics of North American fossil packrat middens: past biodiversity revealed by ancient DNA. *Ecol. Evol.* 10(5):2530–44
101. Moritz C, Agudo R. 2013. The future of species under climate change: resilience or decline? *Science* 341(6145):504–8
102. Mulcahy DG, Macdonald KS III, Brady SG, Meyer C, Barker KB, Coddington J. 2016. Greater than *X* kb: a quantitative assessment of preservation conditions on genomic DNA quality, and a proposed standard for genome-quality DNA. *PeerJ* 4:e2528

103. Murray GGR, Soares AER, Novak BJ, Schaefer NK, Cahill JA, et al. 2017. Natural selection shaped the rise and fall of passenger pigeon genomic diversity. *Science* 358(6365):951–54
104. Nagy ZT. 2010. A hands-on overview of tissue preservation methods for molecular genetic analyses. *Org. Divers. Evol.* 10(1):91–105
105. Natl. Acad. Sci. Eng. Med. 2020. *Biological Collections: Ensuring Critical Research and Education for the 21st Century*. Washington, DC: Natl. Acad. Press
106. Nordling L. 2021. 'The damage is total': Fire rips through historic South African library and plant collection. *Nature* 592(7856):672
107. Novak BJ. 2018. De-extinction. *Genes* 9(11):548
108. Organ CL, Schweitzer MH, Zheng W, Freimark LM, Cantley LC, Asara JM. 2008. Molecular phylogenetics of mastodon and *Tyrannosaurus rex*. *Science* 320(5875):499
109. Orlando L, Ginolhac A, Zhang G, Froese D, Albrechtsen A, et al. 2013. Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499(7456):74–78
110. Parejo M, Wragg D, Henriques D, Charrière J-D, Estonba A. 2020. Digging into the genomic past of Swiss honey bees by whole-genome sequencing museum specimens. *Genome Biol. Evol.* 12(12):2535–51
111. Pasquesi GIM, Adams RH, Card DC, Schield DR, Corbin AB, et al. 2018. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. *Nat. Commun.* 9(1):2774
112. Phifer-Rixey M, Bi K, Ferris KG, Sheehan MJ, Lin D, et al. 2018. The genomic basis of environmental adaptation in house mice. *PLOS Genet.* 14(9):e1007672
113. Poinar HN, Schwarz C, Qi J, Shapiro B, MacPhee RDE, et al. 2006. Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* 311(5759):392–94
114. Poinar HN, Stankiewicz BA. 1999. Protein preservation and DNA retrieval from ancient tissues. *PNAS* 96(15):8426–31
115. Preisser W. 2019. Latitudinal gradients of parasite richness: a review and new insights from helminths of cricetid rodents. *Ecography* 42(7):1315–30
116. Prendini L, Hanner R, DeSalle R. 2002. Obtaining, storing and archiving specimens and tissue samples for use in molecular studies. In *Techniques in Molecular Systematics and Evolution*, ed. R DeSalle, G Giribet, W Wheeler, pp. 176–248. Basel, Switz.: Birkhäuser
117. Pruvost M, Schwarz R, Correia VB, Champlot S, Braguier S, et al. 2007. Freshly excavated fossil bones are best for amplification of ancient DNA. *PNAS* 104(3):739–44
118. Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, et al. 2011. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334(6052):94–98
119. Rasys AM, Park S, Ball RE, Alcala AJ, Lauderdale JD, Menke DB. 2019. CRISPR-Cas9 gene editing in lizards through microinjection of unfertilized oocytes. *Cell Rep.* 28(9):2288–2292.e3
120. Reich D. 2018. *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past*. Oxford, UK: Oxford Univ. Press
121. Riddell EA, Iknayan KJ, Hargrove L, Tremor S, Patton JL, et al. 2021. Exposure to climate change drives stability or collapse of desert mammal and bird communities. *Science* 371(6529):633–36
122. Robertson T, Döring M, Guralnick R, Bloom D, Wieczorek J, et al. 2014. The GBIF Integrated Publishing Toolkit: facilitating the efficient publishing of biodiversity data on the internet. *PLOS ONE* 9(8):e102623
123. Rocha LA, Aleixo A, Allen G, Almeda F, Baldwin CC, et al. 2014. Specimen collection: an essential tool. *Science* 344(6186):814–15
124. Rodriguez D, Becker CG, Pupin NC, Haddad CFB, Zamudio KR. 2014. Long-term endemism of two highly divergent lineages of the amphibian-killing fungus in the Atlantic Forest of Brazil. *Mol. Ecol.* 23(4):774–87
125. Rodríguez Mega E. 2020. Second Brazilian museum fire in two years reignites calls for reform. *Nature* 583(7815):175–76
126. Rouhan G, Dorr LJ, Gautier L, Clerc P, Muller S, Gaudeul M. 2017. The time has come for Natural History Collections to claim co-authorship of research articles. *Taxon* 66(5):1014–16
127. Rowe KC, Rowe KMC, Tingley MW, Koo MS, Patton JL, et al. 2015. Spatially heterogeneous impact of climate change on small mammals of montane California. *Proc. R. Soc. B* 282(1799):20141857

128. Roycroft E, MacDonald AJ, Moritz C, Moussali A, Portela Miguez R, Rowe KC. 2021. Museum genomics reveals the rapid decline and extinction of Australian rodents since European settlement. *PNAS* 118(27):e2021390118
129. Ruane S, Austin CC. 2017. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Mol. Ecol. Resour.* 17(5):1003–8
130. Rubi TL, Knowles LL, Dantzer B. 2020. Museum epigenomics: characterizing cytosine methylation in historic museum specimens. *Mol. Ecol. Resour.* 20(5):1161–70
131. Rubidge EM, Patton JL, Lim M, Burton AC, Brashares JS, Moritz C. 2012. Climate-induced range contraction drives genetic erosion in an alpine mammal. *Nat. Clim. Change* 2(4):285–88
132. Ryder OA, Onuma M. 2018. Viable cell culture banking for biodiversity characterization and conservation. *Annu. Rev. Anim. Biosci.* 6:83–98
133. Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* 364(6435):74–78
134. Schindel DE, Cook JA. 2018. The next generation of natural history collections. *PLOS Biol.* 16(7):e2006125
135. Schmitt CJ, Cook JA, Zamudio KR, Edwards SV. 2019. Museum specimens of terrestrial vertebrates are sensitive indicators of environmental change in the Anthropocene. *Philos. Trans. R. Soc. B* 374(1763):20170387
136. Schubert M, Jónsson H, Chang D, Sarkissian CD, Ermini L, et al. 2014. Prehistoric genomes reveal the genetic foundation and cost of horse domestication. *PNAS* 111(52):E5661–69
137. Schweitzer MH, Zheng W, Organ CL, Avci R, Suo Z, et al. 2009. Biomolecular characterization and protein sequences of the Campanian hadrosaur *B. canadensis*. *Science* 324(5927):626–31
138. Secr. Conven. Biol. Divers. 2011. *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from Their Utilization to the Convention on Biological Diversity*. Montreal: Secr. Conven. Biol. Divers. **https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf**
139. Shapiro B, Drummond AJ, Rambaut A, Wilson MC, Matheus PE, et al. 2004. Rise and fall of the Beringian steppe bison. *Science* 306(5701):1561–65
140. Shapiro B, Hofreiter M. 2014. A paleogenomic perspective on evolution and gene function: new insights from ancient DNA. *Science* 343(6169):1236573
141. Shapiro B, Sibthorpe D, Rambaut A, Austin J, Wragg GM, et al. 2002. Flight of the dodo. *Science* 295(5560):1683
142. Shultz AJ, Adams BJ, Bell KC, Ludt WB, Pauly GB, Vendetti JE. 2021. Natural history collections are critical resources for contemporary and future studies of urban evolution. *Evol. Appl.* 14(1):233–47
143. Shultz AJ, Baker AJ, Hill GE, Nolan PM, Edwards SV. 2016. SNPs across time and space: population genomic signatures of founder events and epizootics in the house finch (*Haemorhous mexicanus*). *Ecol. Evol.* 6(20):7475–89
144. Smith BT, Gehara M, Harvey MG. 2021. The demography of extinction in eastern North American birds. *Proc. R. Soc. B* 2881944:20201945
145. Takei Y, Yun J, Zheng S, Ollikainen N, Pierson N, et al. 2021. Integrated spatial genomics reveals global architecture of single nuclei. *Nature* 590:344–50
146. Termignoni-Garcia F, Louder MIM, Balakrishnan CN, O'Connell L, Edwards SV. 2020. Prospects for sociogenomics in avian cooperative breeding and parental care. *Curr. Zool.* 66(3):293–306
147. Tingley MW, Monahan WB, Beissinger SR, Moritz C. 2009. Birds track their Grinnellian niche through a century of climate change. *PNAS* 106(Suppl. 2):19637–43
148. Toews DPL, Kramer GR, Jones AW, Brennan CL, Cloud BE, et al. 2020. Genomic identification of intergeneric hybrids in New World wood-warblers (Aves: Parulidae). *Biol. J. Linn. Soc.* 131(1):183–91
149. Totoiu CA, Phillips JM, Reese AT, Majumdar S, Girguis PR, et al. 2020. Vortex fluidics-mediated DNA rescue from formalin-fixed museum specimens. *PLOS ONE* 15(1):e0225807
150. van der Valk T, Pečnerová P, Díez-del-Molino D, Bergström A, Oppenheimer J, et al. 2021. Million-year-old DNA unravels the genomic history of mammoths. *Nature* 591:265–69
151. Vershinina AO, Kapp JD, Baryshnikov GF, Shapiro B. 2020. The case of an arctic wild ass highlights the utility of ancient DNA for validating problematic identifications in museum collections. *Mol. Ecol. Resour.* 20(5):1182–90

152. Vo A-TE, Bank MS, Shine JP, Edwards SV. 2011. Temporal increase in organic mercury in an endangered pelagic seabird assessed by century-old museum specimens. *PNAS* 108(18):7466–71
153. Wallis EJ. 2006. Online Zoological Collections of Australian Museums (OZCAM): a national approach to making zoological data available on the web. *Integr. Zool.* 1(2):78–79
154. Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, et al. 2014. Direct evidence of milk consumption from ancient human dental calculus. *Sci. Rep.* 4(1):7104
155. Welker F, Collins MJ, Thomas JA, Wadsley M, Brace S, et al. 2015. Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* 522(7554):81–84
156. Wilson MC, Hills LV, Shapiro B. 2008. Late Pleistocene northward-dispersing *Bison antiquus* from the Bighill Creek Formation, Gallelli Gravel Pit, Alberta, Canada, and the fate of *Bison occidentalis*. *Can. J. Earth Sci.* 45(7):827–59
157. Wisely SM, Ryder OA, Santymire RM, Engelhardt JF, Novak BJ. 2015. A road map for 21st century genetic restoration: gene pool enrichment of the black-footed ferret. *J. Hered.* 106(5):581–92
158. Yan J, Qiu Y, Ribeiro dos Santos AM, Yin Y, Li YE, et al. 2021. Systematic analysis of binding of transcription factors to noncoding variants. *Nature* 591:147–51
159. Yang T, Miller M, Forgacs D, Derr J, Stothard P. 2020. Development of SNP-based genomic tools for the Canadian Bison Industry: parentage verification and subspecies composition. *Front. Genet.* 11:585999
160. Yates TL, Mills JN, Parmenter CA, Ksiazek TG, Parmenter RR, et al. 2002. The ecology and evolutionary history of an emergent disease: hantavirus pulmonary syndrome: Evidence from two El Niño episodes in the American Southwest suggests that El Niño–driven precipitation, the initial catalyst of a trophic cascade that results in a delayed density-dependent rodent response, is sufficient to predict heightened risk for human contraction of hantavirus pulmonary syndrome. *BioScience* 52(11):989–98
161. Zhao M, Wu X, Liao H-T, Liu Y. 2020. Exploring research fronts and topics of Big Data and Artificial Intelligence application for cultural heritage and museum research. *IOP Conf. Ser. Mater. Sci. Eng.* 806(1):012036
162. Zimkus BM, Ford LS. 2014. Best practices for genetic resources associated with natural history collections: recommendations for practical implementation. *Collect. Forum.* 28(1–2):77–112

Annual Review of Genetics

Volume 55, 2021

Contents

Errata

An online log of corrections to *Annual Review of Genetics* articles may be found at http://www.annualreviews.org/errata/genet

Related Articles

From the ***Annual Review of Animal Biosciences***, Volume 9 (2021)

Beyond Antimicrobial Use: A Framework for Prioritizing Antimicrobial Resistance Interventions
Noelle R. Noyes, Ilya B. Slizovskiy, and Randall S. Singer

Epigenetic Regulation of the Nuclear and Mitochondrial Genomes: Involvement in Metabolism, Development, and Disease
Justin C. St. John

Genetics and Evolution of Mammalian Coat Pigmentation
Eduardo Eizirik and Fernanda J. Trindade

The Diversity of Primates: From Biomedicine to Conservation Genomics
Joseph D. Orkin, Lukas F.K. Kuderna, and Tomas Marques-Bonet

From the ***Annual Review of Biochemistry***, Volume 90 (2021)

Repair of DNA Double-Strand Breaks by the Nonhomologous End Joining Pathway
Benjamin M. Stinson and Joseph J. Loparo

Molecular Epigenetics: Chemical Biology Tools Come of Age
John D. Bagert and Tom W. Muir

Optobiochemistry: Genetically Encoded Control of Protein Activity by Light
Jihye Seong and Michael Z. Lin

From the ***Annual Review of Cancer Biology***, Volume 5 (2021)

Zena Werb (1945–2020): Matrix Metalloproteinases, Microenvironments, and Mentoring
Mikala Egeblad

Molecular Heterogeneity and Evolution in Breast Cancer
Jennifer L. Caswell-Jin, Carina Lorenz, and Christina Curtis

The Bidirectional Relationship Between Cancer Epigenetics and Metabolism
Luke T. Izzo, Hayley C. Affronti, and Kathryn E. Wellen

Reeling in the Zebrafish Cancer Models
Alicia M. McConnell, Haley R. Noonan, and Leonard I. Zon

From the ***Annual Review of Cell and Developmental Biology***, Volume 37 (2021)

Beyond Casual Resemblances: Rigorous Frameworks for Comparing Regeneration Across Species
Mansi Srivastava

Components and Mechanisms of Nuclear Mechanotransduction
Philipp Niethammer

Mechanisms of Selective Autophagy
Trond Lamark and Terje Johansen

From the ***Annual Review of Genomics and Human Genetics***, Volume 22 (2021)

A Long, Fulfilling Career in Human Genetics
Haig H. Kazazian Jr.

Transcriptional Regulation by (Super)Enhancers: From Discovery to Mechanisms
Frank Grosveld, Jente van Staalduinen, and Ralph Stadhouders

The Yin and Yang of Histone Marks in Transcription
Paul B. Talbert and Steven Henikoff

Threespine Stickleback: A Model System For Evolutionary Genomics
Kerry Reid, Michael A. Bell, and Krishna R. Veeramah

From the ***Annual Review of Immunology***, Volume 39 (2021)

Epigenetic Remodeling in Innate Immunity and Inflammation
Qian Zhang and Xuetao Cao

Tissue Homeostasis and Inflammation
Matthew L. Meizlish, Ruth A. Franklin, Xu Zhou, and Ruslan Medzhitov

From the ***Annual Review of Microbiology***, Volume 75 (2021)

Hostile Takeover: How Viruses Reprogram Prokaryotic Metabolism
Tyler B. Jacobson, Melanie M. Callaghan, and Daniel Amador-Noguez

Transcriptional Pausing as a Mediator of Bacterial Gene Regulation
Robert Landick

From the ***Annual Review of Phytopathology***, Volume 59 (2021)

One Hundred Years of Hybrid Necrosis: Hybrid Autoimmunity as a Window into the Mechanisms and Evolution of Plant–Pathogen Interactions
Lei Li and Detlef Weigel

From the ***Annual Review of Plant Biology***, Volume 72 (2021)

A Central Role for Genetics in Plant Biology
Maarten Koornneef

Time-Based Systems Biology Approaches to Capture and Model Dynamic Gene Regulatory Networks

Jose M. Alvarez, Matthew D. Brooks, Joseph Swift, and Gloria M. Coruzzi