

Review

Mining museums for historical DNA: advances and challenges in museomics

Christopher J. Raxworthy^{1,*} and Brian Tilston Smith¹

Historical DNA (hDNA), obtained from museum and herbarium specimens, has yielded spectacular new insights into the history of organisms. This includes documenting historical genetic erosion and extinction, discovering species new to science, resolving evolutionary relationships, investigating epigenetic effects, and determining origins of infectious diseases. However, the development of best-practices in isolating, processing, and analyzing hDNA remain under-explored, due to the substantial diversity of specimen preparation types, tissue sources, archival ages, and collecting histories. Thus, for hDNA to reach its full potential, and justify the destructive sampling of the rarest specimens, more experimental work using time-series collections, and the development of improved methods to correct for data asymmetries and biases due to DNA degradation are required.

The advent of hDNA

The potential of natural history museums and herbariums to serve as a vast source of DNA has long been recognized [1–6]. The growing genomic field of **museomics** (see [Glossary](#)) [7], initially yielded short fragments of **historical DNA (hDNA)** that was often from mitochondrial genes, but has since expanded to sequence the entire genome (e.g., [8,9]). Advances in DNA sequencing technology have propelled the field forward, making it logistically simpler and cost-effective to sample across many species' genomes within a single study. However, researchers still face challenges in both producing and analyzing data.

In contrast to hDNA that is fortuitously preserved in historical specimens, **ancient DNA (aDNA)** is naturally preserved and heavily degraded trace DNA in biological material, usually between thousands to a million years old (e.g., [10]; [Box 1](#)). Consequently, DNA obtained from museum **voucher specimens** (almost always collected during the last 200 years) has emerged as a separate subdiscipline, distinct from the study of aDNA. The major significance of this distinction is that optimal methods for sampling, isolating, **bioinformatic** processing, and analyzing hDNA and aDNA, differ due to their contrasting DNA degradation history. These differences are produced because of the natural versus archival preservation conditions, and the different ages of the DNA [11,12].

Why are museum specimens valuable for DNA?

Starting in the 1500s (herbaria) and 1600s (natural history museums), people began to collect, archive, and characterize biodiversity. These efforts continue to the present, although there have been notable declines in collecting of some groups [13] and museums are facing new challenges [14]. Museums and herbaria thus provide an unparalleled spatiotemporal record of species and populations that, on a global scale, now spans the last 200 years. Over the past decade a range of techniques have been developed (or refined) to obtain hDNA from most types of museum specimens ([Table 1](#)). Dried mammal, bird, and reptile skins/mounts [15–17] were among

Highlights

The study of historical DNA (hDNA), utilizing museum specimens, has emerged as a sub-discipline distinct from ancient DNA.

Recent advances in hDNA extraction and sequencing support minimally destructive sampling requests for well-represented specimens, and can now be considered routine/low risk.

Availability of reference genomes, improved mapping techniques, and decreased costs make whole genome resequencing an attractive option for future hDNA studies.

New research is needed to describe and correct for hDNA degradation over decadal-scales, and across diverse tissues and voucher types. This will improve hDNA source selection, and enhance integration with modern DNA. Improved methods are needed for formalin-fixed specimens.

Together, modern and hDNA have the potential to accelerate comparative biological research on a scale not seen since museum collections were first established.

¹Division of Vertebrate Zoology, American Museum of Natural History, 200 Central Park West, New York, NY 10024, USA

*Correspondence: rax@amnh.org (C.J. Raxworthy).



Box 1. Consistent naming of DNA sources

Ancient versus historical versus modern versus archival DNA

There has been confusion in the literature with interchangeable terms being used for various sources of DNA obtained from museum specimens. The most common confusion is between historical DNA (hDNA) and ancient DNA (aDNA), especially in reference to extraction methods and clean room DNA extraction facilities. We follow here other researchers that distinguish the two based on aDNA being naturally preserved in highly degraded trace amounts and >200 years old, and hDNA being derived from specimens archived in collections and almost always <200 years old [1,11]. We also further restrict hDNA to that fortuitously obtained from traditional museum voucher specimens (specimens not intended to serve as a source for DNA). By contrast, tissues (frozen or in preservative) collected as DNA samples, and their DNA extracts (which may have longer archive life spans than currently considered) can be termed as modern DNA. As an example, a recent museomic study [74] successfully obtained DNA from frozen tissues collected in 1982 and more recently; but we consider this to be modern rather than hDNA. Collectively, all sources of DNA stored in museum collections can be termed as **archival DNA**, including: (i) DNA tissues, DNA extracts, and other miscellaneous genetic resources; and (ii) traditional voucher specimens.

Museum formalin-fixed DNA versus formalin-fixed paraffin embedded (FFPE) DNA

DNA exposed to formalin (formaldehyde solution) is often referred to as 'fixed', in reference to the assumption of DNA being stabilized and preserved, similar to the fixation of tissues for histology work, and the fixation of fluid-voucher specimens for museum archives. However, there are two sharply contrasting methods used to fix DNA: (i) formalin-fixed museum specimens (a common fluid preparation type, which became well established in the early 1900s), which typically uses a 10% concentration of saturated formalin solution that is in contact with vertebrate specimens for days to decades [102]; and (ii) clinical FFPE material which uses 2–4% formalin and has much shorter contact times, <20 minutes to a maximum of 48 hours [52,78]. FFPE tissue is also embedded in paraffin, which requires removal during the DNA extraction stage. These differences in formalin concentration and exposure time are known to have profound effects on DNA degradation [48,49]. Researchers may not test alternative formalin-fixation methods during DNA extraction studies [46], and there is a risk that suboptimal methods (e.g., commercial FFPE DNA extraction kits) are applied without modification to museum formalin fixed tissues. For hDNA work on museum formalin-fixed specimens, great care is thus needed to clearly distinguish methodology and advances based on clinical FFPE DNA specimens.

Glossary

Ancient DNA (aDNA): DNA recovered from organic material (e.g., bone, plant matter, environmental samples, and subfossils) that is preserved under natural conditions. This DNA is in trace amounts, heavily degraded, usually >200 years old and often much older.

Archival DNA: all DNA stored in museums, herbaria or other collections as either traditional voucher specimens (historical DNA) or tissue and DNA samples (modern DNA).

Bioinformatics: the field of biology that uses and develops computational methods and software tools for understanding large and complex biological data sets.

Deamination: a common form of degradation of historical (and ancient) DNA that replaces cytosine with uracil.

Depurination: the loss of purine bases from DNA. Hydrolytic depurination is a common form of degradation in ancient DNA that causes fragmentation.

Destructive sampling: the means of collecting a historical DNA sample that requires irreversibly removing and consuming part of the voucher specimen.

Epigenetics: the study of modifications of a cell's DNA, which are reversible modifications that affect gene expression without altering the DNA sequence.

Exogenous (extrinsic) DNA: the contaminating DNA in a sample that does not belong to the target species.

Historical DNA (hDNA): the DNA fortuitously contained in the tissues of traditional voucher specimens that are stored in museums and herbaria. Most voucher specimens in collections are <200 years old.

Hybridization restriction site-associated DNA sequencing

(RADseq): a technique where probes targeting SNPs are designed from sequence fragments produced from RADseq of modern DNA samples.

These probes are then used to target these genomic regions of interest using sequence capture of either modern or historical DNA samples.

Long-read sequencing: a DNA sequencing technique that determines read lengths between 10 000–100 000 base pairs. Also referred to as third-generation sequencing.

Modern DNA: tissues (including blood) and DNA samples (e.g., in alcohol, buffers, frozen, etc.) stored as archives

Table 1. A summary of example museum voucher specimens that have been recently mined for hDNA

hDNA target group	Specimen type	Sources for hDNA	Recent Studies
Infectious diseases	Mammal dry skin Amphibian formalin/alcohol preserved skin	Skin or epithelial swabs	[19,80,82]
Microbial communities	Herbarium dry pressed	Leaf tissue	[18]
Plants	Herbarium dry pressed	Leaf tissue	[7,38,43,44,93,103,104]
Vertebrates	Anthropological material: Headdresses, clothing	Biological material	[105]
Mammals	Skins and hides	Hair, skin, muscle traces, claws	[40,76,87,106]
Mammals	Skeletal	Bone, teeth	[9,87,106]
Birds	Round and flat dry skins	Toe pads, skin	[25,57,64,89,100]
Birds; crabs	Eggs	Piece of egg or powder	[20,107]
Fishes Amphibians Reptiles	Fluid: formalin and/or ethanol preserved	Tissue including liver, muscle	[23,33,35,42,47,58,77,88]
Reptiles	Bone	Carapace, plastron, other bone elements	[34,75,85]
Insects Arachnids	Pinned, dried specimens	Whole limbs, whole specimens, subsamples	[21,22,29,36,61,108–111]
Bryozoa	Dried colonies	Partial specimens	[21,22,29,36,61,108–111]
Arachnids	Fluid: ethanol preserved	Whole or partial specimens	[108]

the first historical specimens to reliably produce hDNA and have since become a frequent sampling target. But recent efforts have obtained DNA from a broader array of museum specimen types and taxonomic groups, including sampling microbial communities from herbarium specimens [18], isolating retroviruses in koala (*Phascolarctos cinereus*) specimens [19], genetic barcoding of eggshells [20], genotyping century-old insect specimens [21], obtaining mitogenomes from 130 year-old dry Bryozoans [22], and collecting genomic loci from formalin-preserved specimens [23]. Of major significance, the use of museum specimens as a source for DNA greatly increases the scale of genetic resources available for research, and consequently, will accelerate comparative studies by substantially increasing taxonomic inclusion, reducing or eliminating field costs, and saving research time.

The recognition that hDNA allows for sampling extinct species (e.g., [24–27]) and extirpated populations (e.g., [28]), as well as improving taxonomic coverage (e.g., [29]), has been borne out in new work (Table 2). Recent hDNA studies have helped to discover and delimit species (e.g., [27,30]), resolve phylogenetic relationships and investigate convergent evolution [31], infer biogeographic history [7], document genetic erosion in populations [9,32], clarify taxonomic confusion with type specimens [33–36], and investigate **epigenetic** effects such as cytosine methylation [37]. Historical specimens have also been used to pinpoint the biogeographic origins of domesticated crops, which often have highly complex anthropogenic histories. For example, hDNA was used to determine the geographic origin of the Coca plant (*Erythroxylum*) in South America [38]. Another exciting application of hDNA is the use of time-series studies to document Anthropocene genetic changes and invasion biology [5,39]. For example, Bi *et al.* [40] reported positive selection on the gene *Alox15* in chipmunk specimens sampled in 1911 and 2012, demonstrating that adaptive responses are recoverable between modern and historical material.

Despite the potential substantial gains in taxonomic sampling, the vast number of museum specimens remain a largely untapped resource for hDNA. The Global Biodiversity Information Facility

to reduce DNA degradation over time. These samples have been primarily collected during the last 40 years.

Museomics: a portmanteau (word blending the sounds and combining the meanings of two other words) of 'museum' and 'genomics' referring to the study of DNA sequences obtained from museum specimens.

RADseq: restriction site associated DNA sequencing, which is a range of similar techniques that use restriction enzymes to fragment DNA. The fragmented products are then sequenced.

Sequence capture: a genomic technique that uses hybridization enrichment with biotinylated RNA or DNA bait sets to target genomic regions of interest.

Short-read sequencing: a DNA sequencing technique that determines read lengths between 200–400 base pairs.

SNP: a nucleotide site that shows base variation (substitutions) between different DNA sequences.

Voucher specimen: a specimen (partial or complete) typically prepared from an individual (but sometimes colony) of animal, fungus, plant, or microbial life that is stored in a museum or herbarium. The voucher specimen is associated with collection data, which together, document and archive the occurrence of a species in space and time.

Table 2. Examples of genomic techniques used for vertebrate, invertebrate, and herbarium specimens

Group	Collection years	Significance	Approach	Refs
Herbarium plant specimens	1835–2009	Obtained genomic loci from low concentrations of DNA from up to 180-year-old specimens	Sequence capture	[43]
<i>Peromyscus</i> Deer Mice	1940–2016	Characterized cytosine methylation in <i>Peromyscus</i> dried skulls	RADseq and bisulfite treatment	[37]
<i>Calyptorhynchus banksii</i> Bird	1896–2014	Showed the efficiency of collecting SNPs using a restriction-site sequencing approach from samples 5–123 years old	Restriction enzyme-based platform	[62]
Snakes	1906–1990	Obtained genomic loci from century-old formalin and ethanol-preserved specimens	Sequence capture	[23]
<i>Gyrinophilus porphyriticus</i> Salamander	1985–1990	Used RADseq to design probes to capture SNPs from allozyme supernatant and formalin preserved specimens	Hybridization RADseq	[47]
Opiliones Spiders	1899–1976	Obtained genomic loci from ethanol-preserved specimens	Sequence capture	[58]
<i>Xylocopa</i> Bees	1894–2013	Obtained genomic loci from century-old pinned specimens	Sequence capture	[29]
<i>Lycaena helle</i> Butterfly	1908–2006	Developed approach where probes were designed from a RAD library, and the probes were used to capture target regions up to century-old museum specimens	Hybridization RADseq	[63]

(GBIF) lists over 84 000 000 organic specimens in both Animalia and Plantae in the world's collections that could yield DNA [GBIF occurrence download (2021) <https://doi.org/10.15468/dl.hq2tjs>], and this excludes the vast number of undigitized collections (e.g., invertebrates). However, most comparative DNA studies still do not include hDNA from voucher specimens, and museum collections continue to be heavily underutilized as a genetic resource.

Recent advances

Extraction methods

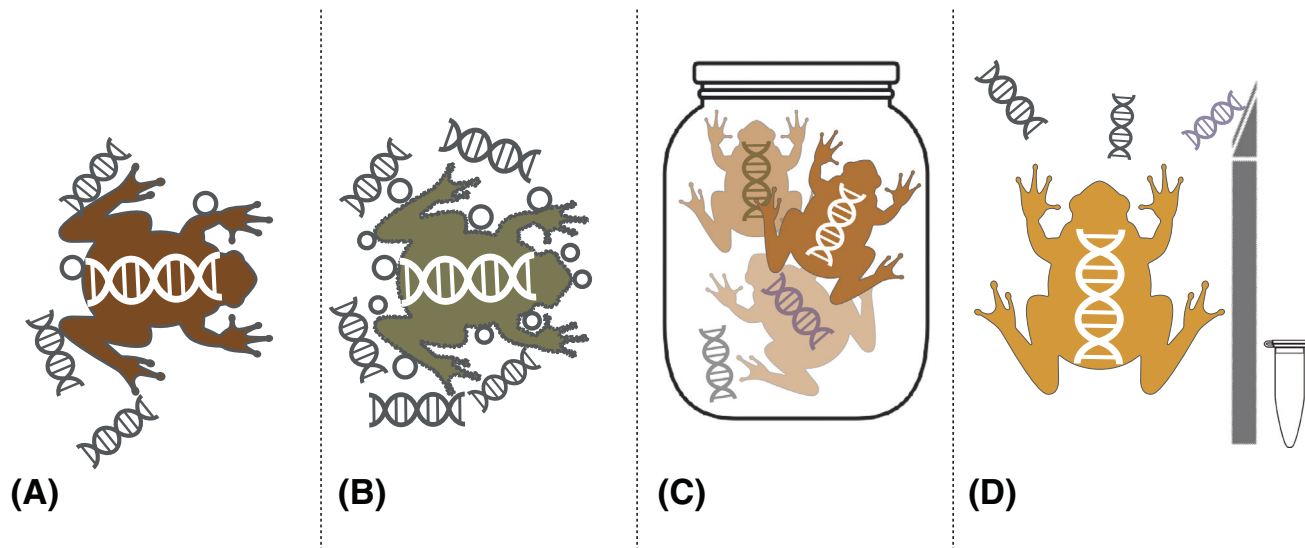
Traditional museum specimens are almost always stored at room temperature as dry or fluid-preserved collections. At these archival room temperatures, high molecular weight DNA is lost over time as the DNA shears to shorter fragments, and DNA quantity also reduces [41]. Although highly desirable, most museums do not have the staff resources to subsample and store hDNA samples in better archival conditions, such as in cryo collections. Another DNA degradation problem occurs with specimens stored in alcohol (typically 70–75% ethanol), which will suffer hydrolytic damage if the alcohol concentration drops due to evaporative loss (a universal problem for fluid collections). Nevertheless, even the oldest historical museum specimens can yield DNA: for example, a 145-year-old alcohol preserved lizard specimen stored at room temperature [42] and herbarium specimens collected from the 1830s (e.g., [43,44]). And more generally, there have been recent methodological advances in extraction success across many specimen types (Table 1). Substantial progress has also been made with developing protocols to reduce **exogenous (extrinsic) DNA** contamination of samples (Figure 1), including adopting aDNA clean room standards such as strict one-way rule of movement, rigorous lab cleaning procedures, and other best-practice protocols for tissue sampling from specimens [45].

For formalin-fixed museum specimens there are special challenges due to DNA cross-linkages with proteins and other strands of DNA, which start to form immediately upon fixation, and that frequently result in complete DNA extraction failure using standard methods [46,47]. DNA damage, including **deamination** and **depurination**, further increases over time (Figure 2) and is accelerated by unbuffered formalin (which is acidic), higher formalin concentrations, and higher temperatures [46,48,49]. Advances have been made with clinical formalin-fixed paraffin embedded (FFPE) material (e.g., [50–52]). However, because formalin fixation methods differ substantially between museum vouchers and clinical FFPE samples (Box 1), these new techniques may not be directly applicable to hDNA. For formalin-fixed hDNA, higher temperatures during buffer incubation [23,42,52,53] and alkali treatments [48,53,54] have both been reported to help break cross-linkages. In addition, several studies have emphasized the importance of removing residual formalin (which may inhibit reagents such as Proteinase K) by washing with GTE buffer [46,47].

The Dabney binding buffer method developed for short-length aDNA [55,56] also offers good prospects for dry and possibly formalin-fixed specimens that do not yield DNA using standard purification methods. These more widely used purification methods include commercial extraction kits, silica bead based methods, and phenol-chloroform extractions (see [11] for a review).

Generating hDNA

The growth of hDNA as a data type has occurred in tandem with advances in DNA sequencing technology. The advent of high-throughput sequencing platforms that require short fragments of DNA have allowed for simultaneously sequencing both more individuals, and more extensively across the genome. Currently, a commonly employed hDNA method is **sequence capture** of targeted loci [23,57–60]. Restriction site associated DNA sequencing approaches [61,62] and **hybridization restriction site-associated DNA sequencing (RADseq)** [63–65] are also effective means of collecting **single nucleotide polymorphisms (SNPs)** from museum

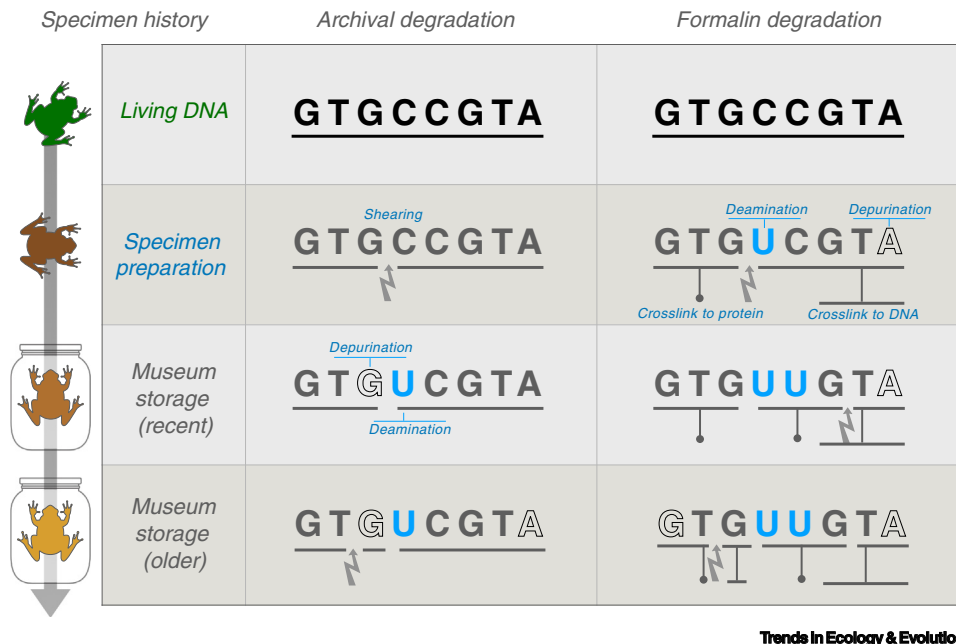


Trends in Ecology & Evolution

Figure 1. Multiple sources of exogenous DNA contamination for a typical museum specimen. (A) DNA from the microbial and parasitic community of the living organism, (B) DNA from post-mortem bacteria during decomposition, (C) DNA from other museum specimens stored in close proximity (e.g., same jar, shelf, or cabinet), and (D) DNA from other sources in the sampling or extraction lab due to poor sampling techniques.

specimens. Researchers have also been sequencing whole genomes from museum specimens for almost a decade [8,66], although this approach was hampered by high costs, lack of access to high performance computing, and implementation problems of bioinformatic pipelines. Whole-genome resequencing for taxa with small to moderately sized genomes is now poised to become the most commonly employed approach for generating hDNA, due to the increased information content and decreased costs. However, for taxa with large and complex genomes, reduced genomic approaches may remain the most viable option.

A continuing limiting factor of genome resequencing is the availability of reference genomes and mapping to those genomes. Whole genomes are being produced at a rapid pace across the Tree of Life but their suitability for hDNA studies vary. The efficiency of mapping reads to a reference genome decreases, the more divergent the two genomes become from one another. Although some hDNA studies may have reference genomes available for the same taxon, or will undertake *de novo* assembly, for most studies the closest reference could be tens of millions of years divergent. The software package TAPAS is specifically designed to improve mapping reads from degraded samples through systematic testing in cases where closely related references are unavailable [67]. However, even when the reference genome and study taxon are the same species, biases can arise depending on whether samples have the reference allele [68]. Another bias during the mapping step is when pseudohaploid sequences are generated by selecting a single nucleotide from a stack of mapped reads [69]. Accounting for varying rates and types of sequencing error likely present in hDNA data during this step can reduce artifactual patterns in downstream analyses [69]. When suitable reference genomes are available, the quality of the reference and the mapping strategy will impact the bioinformatic processing and downstream applications of the data. For example, a comparison of 30 different read mapping approaches for fragmented and degraded DNA, found that obtaining high mapping precision with low levels of reference bias was variable among strategies [70]. Another challenge with **short-read sequencing** of hDNA is that the fragments may be shorter than the sequencing length of Illumina machines, the most popular platform. Thus, it may be more cost effective to run single-read DNA sequencing



Trends in Ecology & Evolution

Figure 2. A stylized depiction of DNA sequence degradation over time. Schematic showing a museum specimen kept in either: (i) typical archival conditions (e.g., a dry specimen or in ethanol); or (ii) formalin solution. DNA shearing, depurination, and deamination occur under both archival conditions, but formalin exposed DNA also suffers from the formation of crosslinks to proteins and other DNA. Under both conditions, DNA concentration also decreases over time.

because typically used paired end reads will be largely overlapping. Obtaining adequate sequencing coverage, minimizing missing data, correcting for DNA degradation, and removing contaminant DNA are all major challenges for genome resequencing of hDNA samples.

The movement of genomics toward **long-read sequencing**, or third generation technologies that yield chromosomal-level genomes, have helped ameliorate data quality issues in genomes produced with earlier technologies. These technologies are not usable for highly degraded hDNA [71], and thus may not be suitable for many genetic samples of non-model organisms (primarily stored in ethanol or tissue buffers, and not flash frozen). However, mapping hDNA reads to chromosomal-level reference genomes will provide insight into how DNA degradation impacts different genomic regions. Consortia have called for sequencing the genomes of all Eurykatotes [72] and for some groups this includes using long-read sequencing [73]. These projects will have to heavily depend on hDNA, because for practical purposes, many species will only be easily sampled by utilizing historical specimens. Despite these obstacles, other recently developed technologies such as linked-read or synthetic long-read sequencing have shown promise for improving genome sequencing from degraded tissue samples [74].

Emerging research areas

Over the next decade, we anticipate the following: (i) expanding the use of formalin-preserved specimens; (ii) verifying the viability of epigenetics and the study of adaptation in response to climate change; and (iii) tracking infectious diseases. These three topics are still highly experimental, and it remains unclear how useful hDNA will ultimately be as a data source.

There are promising signs that formalin-preserved specimens could substantially enhance future hDNA studies (see earlier). The major preparation type for reptiles, amphibians, and fishes is fluid

preservation; and many fluid-preserved specimens (including other groups) were formalin-fixed from the 1900s onwards. To date, there are far fewer studies that have sequenced hDNA from formalin-preserved specimens than dried skins (e.g., [23,47,54,75–77]). For example, Hykin *et al.* [54] obtained the mitochondrial genome from a 30-year-old formalin-fixed lizard (*Anolis carolinensis*). Ruane and Austin [23] showed the utility of formalin-preserved snakes by generating thousands of short length ultraconserved element (UCE) loci from specimens up to 100 years old. O'Connell *et al.* [47] obtained sequence data from 35-year-old salamander specimens fixed in 10% formalin for days to months, and obtained DNA fragment distribution peaks of 400 base pairs for liver and 185 for muscle tissue. Current trends suggest techniques (and success rates) are continuing to improve, as a result of using combinations of long washes, higher temperatures, and alkali treatments to remove formalin and break cross-linkages.

The majority of hDNA studies focus on evolutionary relationships and patterns of neutral genetic diversity, yet research on adaptation is poised for growth because genomes from hDNA can be annotated using increasingly available reference genomes. A promising area of research uses epigenetic techniques, which have the transformative potential of showing how gene regulation has changed during rapid climate change over the past century [78]. For example, Rubi *et al.* [37] used a **RADseq** technique and bisulfite treatment to estimate methylation in deer mice (*Peromyscus* spp) specimens over 70 years old. The overall viability of epigenetic techniques for museum specimens requires additional vetting because the oldest specimen in Rubi *et al.* [37] had the lowest methylation and showed signs of hydrolytic deamination. Nevertheless, further study of allelic diversity and mutational patterns among historical and modern populations will undoubtedly yield further insights into how selection has shaped populations over the last century.

Calls have also been made to expand the use of hDNA to track the origins and spread of infectious diseases [5,79,80], particularly given the uncertainty surrounding the source of the COVID-19 outbreak [37,81]. For example, white-nose syndrome caused by a fungus (*Pseudogymnoascus destructans*) has recently devastated Nearctic bat populations. hDNA techniques found *P. destructans* DNA in a century-old bat specimen from France, but absent from North American specimens (1861–1971), thus confirming an Old World origin [82]. Similarly, skin swabs from frog museum specimens have been used to track the historical origin of the pathogenic chytrid fungus Bd (*Batrachochytrium dendrobatidis*) [80]. In this regard alone, museum collections provide a remarkable record of historical infectious diseases which otherwise, may be intractable using any other evidence.

Unresolved challenges

Almost since the inception of isolating hDNA from museum specimens, researchers have experimented with alternative approaches and specimen types for acquiring DNA [16,83–86]. This practice has continued through contemporary research that includes evaluating different types and parts of the specimen, testing commercial kits, and modifying ‘home-brew’ approaches [87–89]. One of the general findings is that DNA concentration varies among parts of the specimen and preservation techniques [59,87,89,90], tissue types [54], and sources of genomic DNA [47]. Time-series studies, which have investigated DNA degradation in specimens of varying ages, have shown that the molecular weight and concentration of DNA decreases with time [29,57,91,92]. Observations of allelic dropout, and fewer and shorter loci are clear indications of DNA degradation in older specimens. However, although general trends have been reported, this process still remains poorly understood for most specimen preparation types, and typical archival conditions. Additional studies, using diverse museum vouchers, that quantify the rates of DNA shearing, depurination, and deamination across decadal-scales are needed,

to provide a more nuanced view on the temporal scale that biases arise in hDNA datasets (Figure 2).

A further complication of hDNA concerns the unique collection history of individual specimens, and the resulting high variability in hDNA quality. Collection history includes collection methods, post-mortem decomposition time, field preparation methods, field storage and environmental conditions, archival storage conditions, and archival time duration. All these factors influence the preservation of hDNA, yet are often incompletely documented, especially concerning the preparation history in the field. For example, most fluid vertebrate specimens collected in the 1900s were field fixed in either ethanol or formalin [76,85], yet the actual used fixative is typically undocumented. Post-mortem decomposition time also varies widely (e.g., road-kill versus fresh specimen) and has significant impacts on liver DNA quality which, unlike muscle, is vulnerable to rapid degradation from elevated enzymatic action [41]. Consequently, improved documentation of individual specimen hDNA is desirable. This might be possible to estimate by quantifying DNA content in a specific tissue type and/or incorporating archival and specimen condition metrics (e.g., specimen age and physical condition, and environmental conditions). However, this requires extensive experimental investigation and validation, collaborations between museums, and ideally should be coordinated by consortia or specialist working groups.

It is likely that DNA extraction techniques can be further optimized, especially for particular specimen or tissue types (e.g., formalin-fixed specimens). Frequently, researchers rely on commercial kits that were designed for fresh tissue and blood. DNA concentration has been shown to be higher using phenol-chloroform extraction over column-based extractions [89] which have a size selection step that removes DNA fragments of a particular size. Many laboratories using hDNA rely on 'home-brew' approaches that were often originally optimized to a narrow set of conditions, if at all. One of the outstanding challenges is thus automating DNA extraction protocols across hundreds or even thousands of individuals. Folk *et al.* [93] developed a high-throughput pipeline for building large-scale phylogenies from herbarium collections that includes plate-based extractions and a community science option for metadata aggregation. Automating additional DNA extraction pipelines will allow researchers to take advantage of the massively-parallel sequencing capabilities of short-read sequencing and resolve data generation bottlenecks. Further testing and development of extraction techniques for a broad range of specimens, tissue types, and organisms under different storage conditions will also provide clearer methodological choices.

Recent work has also explored methods to enhance each step of generating higher quality hDNA. A recently validated single-tube library preparation method increased library complexity and decreased loss of DNA template, which improved mapping and processing times [94]. Additional quality control measures include enzymatic repair prior to sequencing [2], removing exogenous DNA by mapping to contaminant genomes [95], rescaling base quality scores based on their probability of being damaged [96], and detecting and filtering out base misincorporation [2]. A variant caller specifically designed for degraded DNA was shown to outperform the most commonly employed variant callers used in genomic studies [97]. Although these quality control measures will remove unwanted noise in hDNA, an asymmetry between the information content and the amount of missing data between historical and modern samples will often persist.

Studies that combined historical and **modern DNA** have revealed a number of biases, including hDNA samples that cannot be accurately placed in phylogenies estimated with the multispecies coalescent [98], clustering based on sample type that is independent of ancestry [99], highly variable branch lengths in phylogenetic trees [47, 100], and overestimates of homozygosity [62]. DNA

sourced from both allozyme homogenates and formalin-fixed tissues (using the same individual specimen) have also shown biases in some analyses: RADseq replicates did not cluster in multivariate space with corresponding capture-based replicates, in contrast to phylogenetic analyses where all replicates clustered together [47]. All these results emphasize the importance of evaluating the robustness of genetic patterns in hDNA by filtering datasets to make them more comparable, subsampling fewer loci that have more comparable missing data, more directly targeting loci/SNPs of interest with probes, or dedicating more sequencing effort towards historical samples.

The curator's conundrum

Curators and collection staff of natural history museums and herbaria are increasingly receiving requests for **destructive sampling** of specimens for hDNA. The conundrum they face is whether to sample irreplaceable material for studies that currently yield desirable but unpredictable results, or stymie research with the hopes that future technologies will produce more reliable and complete genomic data. These decisions are confounded by the current increase of hDNA sampling requests (which are labor intensive to fulfill) coinciding with the general understaffing of museums [101].

There is an urgent need for the research community to adopt museomic best practices, informed on new rigorous experimental work, to justify destructive sampling [78]. And comparative studies that report hDNA failure rates across specimen types and taxa (e.g., [23,76,85]), will be especially useful for the development of best-curation practices regarding destructive sampling. For example, Lalonde and Marcus [21] report extraction and sequencing failure rates for a high-throughput genotyping protocol.

Based upon the recent studies and techniques we review here (Tables 1 and 2), hDNA sampling requests of well-represented specimens that are minimally destructive (e.g., removing minimal tissue from nonunique morphological features) should now be considered routine and low risk. Prioritizing projects that sequence broadly across the genome, or produce sequence data that can be incorporated into other studies versus methods that generate study-specific data (e.g., SNPs) may also be appropriate. However, museums will almost always evaluate requests for destructive sampling on a case-by-case basis, and certainly for rare, unique, or delicate specimens, greater scrutiny should continue, and is both justified and prudent.

Concluding remarks

Recent advances in sequencing methods can now be routinely applied to many traditional museum specimens. Consequently, hDNA is now playing an important role in diverse topics across biology, and its utility will accelerate comparative biological studies on a scale not seen since museum collections were first established. However, unresolved issues still require further targeted research (see Outstanding questions). In particular, a better understanding of the rate of hDNA degradation over decadal-scales is needed to develop a predictive framework for determining the quality of hDNA from individual voucher specimens and tissue sources. Although the field and archive collection history of most specimens are often incompletely documented, we expect common DNA degradation patterns will emerge, which can be used to facilitate the future development of optimal DNA extraction and bioinformatic processing. These advances offer the potential of developing best-practice destructive sampling protocols for rare specimens, and ultimately they will allow hDNA and traditional museum collections to meet their full and incredible research potential.

Acknowledgments

Development of this paper was partly funded by the National Science Foundation (DBI 2029955 to C.J.R. and B.T.S.); Julia Zichello produced the figures and provided editorial input; Kaiya Provost, Lucas Moreira, and Greg Thom provided feedback.

Outstanding questions

What hDNA extraction protocols are most suitable for automation to address large-scale sampling bottlenecks? What tissues and specimens (type and condition) are the best sources of hDNA? Can DNA extraction success be further improved for formalin-fixed specimens? Well-designed experimental studies are needed to explore these questions. New bioinformatic tools are also required to improve mapping, correct for biases and errors, and account for asymmetric information content between historical and modern samples.

How does hDNA degrade across the genome in museum and herbarium specimens during initial preparation, and over decades to centuries of archival storage time? Is it possible to predict the quality of hDNA from individual museum vouchers using known collection parameters and specimen condition? The diversity of preparation types and sources of hDNA, and the variation in field and archival histories, make these complex problems to investigate. The exact nature of DNA degradation is also not well understood for preservatives used in museum specimens such as formalin, alcohol, and skin-tanning chemicals.

Can destructive sampling best-practices become broadly adopted by herbaria and museums that will support both wise collection stewardship and research progress? Progress in this area will depend on developing more reliable hDNA extraction protocols and developing working groups within museum consortia to coordinate this initiative.

Will greater support be provided to natural history collections and herbaria to expand the availability of hDNA, to address the many emerging research applications? Demand for hDNA sampling is expected to soar over the coming decade, yet many museums do not have the current staff and lab resources to facilitate this expected surge in requests.

Declaration of interests

No interests to declare.

References

- Wandeler, P. *et al.* (2007) Back to the future: museum specimens in population genetics. *Trends Ecol. Evol.* 22, 634–642
- Bi, K. *et al.* (2013) Unlocking the vault: next-generation museum population genomics. *Mol. Ecol.* 22, 6018–6032
- Holmes, M.W. *et al.* (2016) Natural history collections as windows on evolutionary processes. *Mol. Ecol.* 25, 864–881
- Yeates, D.K. *et al.* (2016) Museums are biobanks: unlocking the genetic potential of the three billion specimens in the world's biological collections. *Curr. Opin. Insect Sci.* 18, 83–88
- Bleker, V.C. and Martin, M.D. (2018) Implications and future prospects for evolutionary analyses of DNA in historical herbarium collections. *Bot. Lett.* 165, 409–418
- Lopez, L. *et al.* (2020) Genomics of natural history collections for understanding evolution in the wild. *Mol. Ecol. Resour.* 20, 1153–1160
- Schaefer, H. *et al.* (2009) Gourds afloat: a dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous oversea dispersal events. *Proc. Biol. Sci.* 276, 843–851
- Hung, C.-M. *et al.* (2014) Drastic population fluctuations explain the rapid extinction of the passenger pigeon. *Proc. Natl. Acad. Sci. U. S. A.* 111, 10636–10641
- van der Valk, T. *et al.* (2019) Historical genomes reveal the genomic consequences of recent population decline in Eastern gorillas. *Curr. Biol.* 29, 165–170.e6
- van der Valk, T. *et al.* (2021) Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* 591, 265–269
- Billerman, S.M. and Walsh, J. (2019) Historical DNA as a tool to address key questions in avian biology and evolution: a review of methods, challenges, applications, and future directions. *Mol. Ecol. Resour.* 19, 1115–1130
- Mitchell, K.J. and Rawlence, N.J. (2021) Examining natural history through the lens of palaeogenomics. *Trends Ecol. Evol.* 36, 258–267
- Malaney, J.L. and Cook, J.A. (2018) A perfect storm for mammalogy: declining sample availability in a period of rapid environmental degradation. *J. Mammal.* 99, 773–788
- Watanabe, M.E. (2015) The Nagoya Protocol on access and benefit sharing: international treaty poses challenges for biological collections. *Bioscience* 65, 543–550
- Higuchi, R. *et al.* (1984) DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312, 282–284
- Mundy, N.I. *et al.* (1997) Skin from feet of museum specimens as a non-destructive source of DNA for avian genotyping. *Auk* 114, 126–129
- Caccone, A. *et al.* (1999) Origin and evolutionary relationships of giant Galápagos tortoises. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13223–13228
- Bleker, V.C. *et al.* (2020) Metagenomic analysis of historical herbarium specimens reveals a postmortem microbial community. *Mol. Ecol. Resour.* 20, 1206–1219
- Ávila-Arcos, M.C. *et al.* (2013) One hundred twenty years of koala retrovirus evolution determined from museum skins. *Mol. Biol. Evol.* 30, 299–304
- Greal, A. *et al.* (2021) Genetic barcoding of museum eggshell improves data integrity of avian biological collections. *Sci. Rep.* 11, 1605
- Lalonde, M.M.L. and Marcus, J.M. (2020) How old can we go? Evaluating the age limit for effective DNA recovery from historical insect specimens. *Syst. Entomol.* 45, 505–515
- Orr, R.J.S. *et al.* (2021) A molecular phylogeny of historical and contemporary specimens of an under-studied micro-invertebrate group. *Ecol. Evol.* 11, 309–320
- Ruane, S. and Austin, C.C. (2017) Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Mol. Ecol. Resour.* 17, 1003–1008
- Feigin, C.Y. *et al.* (2018) Genome of the Tasmanian tiger provides insights into the evolution and demography of an extinct marsupial carnivore. *Nat. Ecol. Evol.* 2, 182–192
- Smith, B.T. *et al.* (2021) The demography of extinction in eastern North American birds. *Proc. Biol. Sci.* 288, 20201945
- Roycroft, E. *et al.* (2021) Museum genomics reveals the rapid decline and extinction of Australian rodents since European settlement. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2021390118
- Grewe, F. *et al.* (2021) Museum genomics reveals the Xerces blue butterfly (*Glaucopsyche xerces*) was a distinct species driven to extinction. *Biol. Lett.* 17, 20210123
- Shepherd, L.D. and Lambert, D.M. (2008) Ancient DNA and conservation: lessons from the endangered kiwi of New Zealand. *Mol. Ecol.* 17, 2174–2184
- Blaimer, B.B. *et al.* (2016) Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLoS One* 11, e0161531
- Lyra, M.L. *et al.* (2020) High-throughput DNA sequencing of museum specimens sheds light on the long-missing species of the *Bokermannohyla claresignata* group (Anura: Hylidae: Cophomantini). *Zool. J. Linnean Soc.* 190, 1235–1255
- Lamichhane, S. *et al.* (2019) Integrating natural history collections and comparative genomics to study the genetic architecture of convergent evolution. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374, 20180248
- Dussex, N. *et al.* (2019) Complete genomes of two extinct New Zealand passerines show responses to climate fluctuations but no evidence for genomic erosion prior to extinction. *Biol. Lett.* 15, 20190491
- Silva, P.C. *et al.* (2017) Using ancient DNA to unravel taxonomic puzzles: the identity of *Deuterodon pedri* (Ostariophysi: Characidae). *Neotrop. Ichthyol.* 15, e160141
- Kehlmaier, C. *et al.* (2019) Mitogenomics of historical type specimens of Australasian turtles: clarification of taxonomic confusion and old mitochondrial introgression. *Sci. Rep.* 9, 5841
- Kehlmaier, C. *et al.* (2020) The enigmatic Crimean green lizard (*Lacerta viridis magnifica*) is extinct but not valid: mitogenomics of a 120-year-old museum specimen reveals historical introduction. *J. Zool. Syst. Evol. Res.* 58, 303–307
- Cong, Q. *et al.* (2021) Genomics reveals the origins of historical specimens. *Mol. Biol. Evol.* 38, 2166–2176
- Rubi, T.L. *et al.* (2020) Museum epigenomics: characterizing cytosine methylation in historic museum specimens. *Mol. Ecol. Resour.* 20, 1161–1170
- White, D.M. *et al.* (2021) The origins of coca: museum genomics reveals multiple independent domestications from progenitor *Erythroxylum gracilipes*. *Syst. Biol.* 70, 1–13
- Schmitt, C.J. *et al.* (2018) Museum specimens of terrestrial vertebrates are sensitive indicators of environmental change in the Anthropocene. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 374, 20170387
- Bi, K. *et al.* (2019) Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLoS Genet.* 15, e1008119
- Camacho-Sanchez, M. *et al.* (2013) Preservation of RNA and DNA from mammal samples under field conditions. *Mol. Ecol. Resour.* 13, 663–673
- McGuire, J.A. *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
- Hart, M.L. *et al.* (2016) Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* 65, 1081–1092
- Gutaker, R.M. *et al.* (2017) Extraction of ultrashort DNA molecules from herbarium specimens. *Biotechniques* 62, 76–79
- Fulton, T.L. and Shapiro, B. (2019) Setting up an ancient DNA laboratory. *Methods Mol. Biol.* 1963, 1–13
- Zimmermann, J. *et al.* (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. *Front. Zool.* 5, 18
- O'Connell, K.A. *et al.* (2021) Genomic library preparation and hybridization capture of formalin-fixed tissues and allozyme

- supernatant for population genomics and considerations for combining capture- and RADseq-based single nucleotide polymorphism data sets. *Mol. Ecol. Resour.* <https://doi.org/10.1111/1755-0998.13481> Published online July 30, 2021
48. Campos, P.F. and Gilbert, T.M.P. (2012) DNA extraction from formalin-fixed material. *Methods Mol. Biol.* 840, 81–85
 49. Do, H. and Dobrovic, A. (2015) Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin. Chem.* 61, 64–71
 50. Coombs, N.J. et al. (1999) Optimisation of DNA and RNA extraction from archival formalin-fixed tissue. *Nucleic Acids Res.* 27, e12
 51. Haile, S. et al. (2019) Sources of erroneous sequences and artificial chimeric reads in next generation sequencing of genomic DNA from formalin-fixed paraffin-embedded samples. *Nucleic Acids Res.* 47, e12
 52. Robbe, P. et al. (2018) Clinical whole-genome sequencing from routine formalin-fixed, paraffin-embedded specimens: pilot study for the 100 000 Genomes Project. *Genet. Med.* 20, 1196–1205
 53. Gilbert, M.T.P. et al. (2007) The isolation of nucleic acids from fixed, paraffin-embedded tissues—which methods are useful when? *PLoS One* 2, e537
 54. Hykin, S.M. et al. (2015) Fixing formalin: a method to recover genomic-scale DNA sequence data from formalin-fixed museum specimens using high-throughput sequencing. *PLoS One* 10, e0141579
 55. Dabney, J. and Meyer, M. (2019) Extraction of highly degraded DNA from ancient bones and teeth. *Methods Mol. Biol.* 1963, 25–29
 56. Rohland, N. et al. (2018) Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing. *Nat. Protoc.* 13, 2447–2461
 57. McCormack, J.E. et al. (2016) Sequence capture of ultraconserved elements from bird museum specimens. *Mol. Ecol. Resour.* 16, 1189–1203
 58. Derkarabetian, S. et al. (2019) Sequence capture phylogenomics of historical ethanol-preserved museum specimens: unlocking the rest of the vault. *Mol. Ecol. Resour.* 19, 1531–1544
 59. Forrest, L.L. et al. (2019) The limits of Hyb-Seq for herbarium specimens: impact of preservation techniques. *Front. Ecol. Evol.* 7, 439
 60. Trevisan, B. et al. (2019) Genome skimming is a low-cost and robust strategy to assemble complete mitochondrial genomes from ethanol preserved specimens in biodiversity studies. *PeerJ* 7, e7543
 61. Tin, M.M.-Y. et al. (2014) Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. *PLoS One* 9, e96793
 62. Ewart, K.M. et al. (2019) Museum specimens provide reliable SNP data for population genomic analysis of a widely distributed but threatened cockatoo species. *Mol. Ecol. Resour.* 19, 1578–1592
 63. Suchan, T. et al. (2016) Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS One* 11, e0151651
 64. Linck, E.B. et al. (2017) Evaluating hybridization capture with RAD probes as a tool for museum genomics with historical bird specimens. *Ecol. Evol.* 7, 4755–4767
 65. Lang, P.L.M. et al. (2020) Hybridization ddRAD-sequencing for population genomics of nonmodel plants using highly degraded historical specimen DNA. *Mol. Ecol. Resour.* 20, 1228–1247
 66. Staats, M. et al. (2013) Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. *PLoS One* 8, e69189
 67. Taron, U.H. et al. (2018) Testing of alignment parameters for ancient samples: evaluating and optimizing mapping parameters for ancient samples using the TAPAS tool. *Genes* 9, 157
 68. Günther, T. and Nettelblad, C. (2019) The presence and impact of reference bias on population genomic studies of prehistoric human populations. *PLoS Genet.* 15, e1008302
 69. Barlow, A. et al. (2020) Consensify: a method for generating pseudohaploid genome sequences from palaeogenomic datasets with reduced error rates. *Genes* 11, 50
 70. Oliva, A. et al. (2021) Systematic benchmark of ancient DNA read mapping. *Brief. Bioinform.*, bbab076
 71. van Dijk, E.L. et al. (2018) The third revolution in sequencing technology. *Trends Genet.* 34, 666–681
 72. Lewin, H.A. et al. (2018) Earth BioGenome Project: sequencing life for the future of life. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4325–4333
 73. Rhie, A. et al. (2021) Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592, 737–746
 74. Colella, J.P. et al. (2020) A linked-read approach to museomics: higher quality *de novo* genome assemblies from degraded tissues. *Mol. Ecol. Resour.* 20, 856–870
 75. Kearney, M. and Stuart, B.L. (2004) Repeated evolution of limblessness and digging heads in worm lizards revealed by DNA from old bones. *Proc. Biol. Sci.* 271, 1677–1683
 76. O'Toole, B. et al. (2021) Reconstructing the genomic diversity of a widespread sub-Saharan bat (*Pteropodidae: Eidolon helvum*) using archival museum collections. *Acta* 22, 227–241
 77. Appleyard, S.A. et al. (2021) Assessing DNA for fish identifications from reference collections: the good, bad, and ugly shed light on formalin fixation and sequencing approaches. *J. Fish Biol.* 98, 1421–1432
 78. Hahn, E.E. et al. (2020) Museum epigenomics: charting the future by unlocking the past. *Trends Ecol. Evol.* 35, 295–300
 79. DiEuliis, D. et al. (2016) Opinion: specimen collections should have a much bigger role in infectious disease research and response. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4–7
 80. Byrne, A.Q. et al. (2019) Cryptic diversity of a widespread global pathogen reveals expanded threats to amphibian conservation. *Proc. Natl. Acad. Sci. U. S. A.* 116, 20382–20387
 81. Thompson, C.W. et al. (2021) Preserve a voucher specimen! The critical need for integrating natural history collections in infectious disease studies. *MBio* 12, e02698–e02720
 82. Campana, M.G. et al. (2017) White-nose syndrome fungus in a 1918 bat specimen from France. *Emerg. Infect. Dis.* 23, 1611–1612
 83. Ellegren, H. (1994) Genomic DNA from museum bird feathers. In *Ancient DNA: Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical, and Forensic Specimens* (Herrmann, B. and Hummel, S., eds), pp. 211–217. Springer, New York
 84. Wisely, S.M. et al. (2004) A technique for sampling ancient DNA that minimizes damage to museum specimens. *Conserv. Genet.* 5, 105–107
 85. Stuart, B.L. et al. (2006) Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. *Syst. Biodivers.* 4, 133–136
 86. Morin, P.A. et al. (2007) Comparative mitochondrial and nuclear quantitative PCR of historical marine mammal tissue, bone, baleen, and tooth samples. *Mol. Ecol. Notes* 7, 404–411
 87. McDonough, M.M. et al. (2018) Performance of commonly requested destructive museum samples for mammalian genomic studies. *J. Mammal.* 99, 789–802
 88. Silva, P.C. et al. (2019) Comparison and optimization for DNA extraction of archived fish specimens. *MethodsX* 6, 1433–1442
 89. Tsai, W.L.E. et al. (2020) More than skin and bones: comparing extraction methods and alternative sources of DNA from avian museum specimens. *Mol. Ecol. Resour.* 20, 1220–1227
 90. Casas-Marce, M. et al. (2010) Searching for DNA in museum specimens: a comparison of sources in a mammal species. *Mol. Ecol. Resour.* 10, 502–507
 91. Wandeler, P. et al. (2003) Patterns of nuclear DNA degeneration over time—a case study in historic teeth samples. *Mol. Ecol.* 12, 1087–1093
 92. McGaughan, A. (2020) Effects of sample age on data quality from targeted sequencing of museum specimens: what are we capturing in time? *BMC Genomics* 21, 188
 93. Folk, R.A. et al. (2021) High-throughput methods for efficiently building massive phylogenies from natural history collections. *Appl. Plant Sci.* 9, e11410
 94. Caroe, C. et al. (2018) Single-tube library preparation for degraded DNA. *Methods Ecol. Evol.* 9, 410–419
 95. Wingett, S.W. and Andrews, S. (2018) FastQ Screen: a tool for multi-genome mapping and quality control. *F1000Res.* 7, 1338
 96. Jónsson, H. et al. (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684

97. Zhou, B. *et al.* (2017) AntCaller: an accurate variant caller incorporating ancient DNA damage. *Mol. Gen. Genomics*. 292, 1419–1430
98. Hosner, P.A. *et al.* (2016) Avoiding missing data biases in phylogenomic inference: an empirical study in the Landfowl (Aves: Galliformes). *Mol. Biol. Evol.* 33, 1110–1125
99. Smith, B.T. *et al.* (2020) Uneven missing data skew phylogenomic relationships within the lories and lorikeets. *Genome Biol. Evol.* 12, 1131–1147
100. Harvey, M.G. *et al.* (2020) The evolution of a tropical biodiversity hotspot. *Science* 370, 1343–1348
101. Pennisi, E. (2020) Shuttered natural history museums fight for survival. *Science* 368, 1042–1043
102. Simmons, J.E. (2014) *Fluid Preservation: A Comprehensive Reference*, Rowman & Littlefield
103. Silva, C. *et al.* (2017) Museomics resolve the systematics of an endangered grass lineage endemic to north-western Madagascar. *Ann. Bot.* 119, 339–351
104. Slimp, M. *et al.* (2021) On the potential of Angiosperms353 for population genomic studies. *Appl. Plant Sci.* 9, e11419
105. Hartnup, K. *et al.* (2011) Ancient DNA recovers the origins of Māori feather cloaks. *Mol. Biol. Evol.* 28, 2741–2750
106. Hawkins, M.T.R. *et al.* (2016) Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. *BMC Evol. Biol.* 16, 80
107. van der Meij, S.E.T. and Nieman, A.M. (2016) Old and new DNA unweave the phylogenetic position of the eastern Atlantic gall crab *Detocarcinus balssi* (Monod, 1956) (Decapoda: Cryptochiridae). *J. Zool. Syst. Evol. Res.* 54, 189–196
108. Wood, H.M. *et al.* (2018) Next-generation museum genomics: phylogenetic relationships among palpimanoid spiders using sequence capture techniques (Araneae: Palpimanoidea). *Mol. Phylogenet. Evol.* 127, 907–918
109. Gauthier, J. *et al.* (2020) Museomics identifies genetic erosion in two butterfly species across the 20th century in Finland. *Mol. Ecol. Resour.* 20, 1191–1205
110. Jin, M. *et al.* (2020) Museomics reveals extensive cryptic diversity of Australian prionine longhorn beetles with implications for their classification and conservation. *Syst. Entomol.* 45, 745–770
111. Mayer, C. *et al.* (2021) Adding leaves to the Lepidoptera tree: capturing hundreds of nuclear genes from old museum specimens. *Syst. Entomol.* 46, 649–671