ARPHA Preprints

PENSOFT

PREPRINT

Author-formatted, not peer-reviewed document posted on 09/06/2025

DOI: https://doi.org/10.3897/arphapreprints.e161346

Museomics, the *Extended Specimen* and *Collectomics* – how to frame and name the diversity of information linked to specimens in natural history collections?

Martin Kapun, Martin Schwentner, Elisabeth Haring, Nesrine Akkari, Andreas Kroh, Luise Kruckenhauser, Anja Palandacic, Katrin Vohland

ARPHA Preprints *Author-formatted, not peer-reviewed document posted on 09/06/2025*. DOI: https://doi.org/10.3897/arphapreprints.e161346

Natural History Collections and Museomics (pensoft.net) [NHMWMuseomics]

Museomics, the *Extended Specimen* and *Collectomics* – how to frame and name the diversity of information linked to specimens in natural history collections?

Martin Kapun*, Martin Schwentner*, Elisabeth Haring*, Nesrine Akkari, Andreas Kroh, Luise Kruckenhauser, Anja Palandačić, Katrin Vohland§

* equal contribution

§ corresponding author: katrin.vohland@nhm.at

ORCIDS

Martin Kapun: 0000-0002-3810-0504
Martin Schwentner: 0000-0002-1373-456X
Elisabeth Haring: 0000-0002-5411-1879
Nesrine Akkari: 0000-0001-5019-4833
Andreas Kroh: 0000-0002-8566-8848
Luise Kruckenhauser: 0000-0001-8708-4347
Anja Palandačić: 0000-0002-4555-5240
Katrin Vohland: 0000-0002-7214-7015

Keywords

museum collections; collection genomics; persistent identifier; historical DNA; archival DNA

Abstract

263 words

In light of recent technological advances – in fields such as computer science, imaging technologies, and molecular analysis methods – the possibilities for studying biological museum specimens and linking information across disciplines have expanded exponentially in recent years. In particular, scientific disciplines operating on the scale of so-called -omics technologies, originally developed in molecular biology, have found their way into the museum realm. This development has led to a proliferation of neologisms aimed at describing comprehensive approaches for analysing museum specimens and collections with the help of large-scale datasets. Among these terms, “museomics” has become a frequently used, albeit rather loosely defined, synonym for a wide range of large-scale, integrative scientific approaches that are centered on or incorporate museum specimens. In this article, we aim to provide an overview of the most commonly used key terms, examine their conceptual interrelations, and present a literature-based survey of their usage. We advocate for using three terms at different hierarchical levels: (i) *museomics* for analytical approaches using biomolecules on museum specimen; (ii) the *extended specimen* concept linking voucher specimens with diverse physical, digital, and contextual information; and (iii) *collectomics* linking the data of the extended specimen with quantitative datasets from other domains of research. A key prerequisite for this linkage is the unambiguous assignment of data to the corresponding specimen using a Persistent Identifier (PID). Finally, we conclude the paper by reflecting on the current approaches and initiatives employed at the Natural History Museum Vienna, emphasising the importance of adopting a museum-wide strategy for developing extended specimen datasets using standardised data formats in accordance with Open Science principles.

Introduction

Natural history collections provide an extraordinary and unparalleled record of global biodiversity. Collectively, natural history museums preserve billions of specimens from all geographic regions of the planet, most of which were collected over the past two to three centuries (Johnson et al. 2023, Sigwart et al. 2025). When fossils and geological objects are included, these collections reflect hundreds, thousands, or even billions of years of evolutionary history. They offer a unique opportunity to study historical specimens – some collected decades or centuries ago – serving as snapshots of natural ecosystems before the onset of significant human impact (Card et al. 2021, Davis and Knapp 2024). In addition,

natural history collections contain rare or irreplaceable biological specimens that cannot, or should not, be collected again – either because they originate from inaccessible regions or because the species are now rare, threatened, or extinct. Specimens, whether they are single organisms (e.g. one pinned insect), or part of an individual (tissue, or even a footprint in mammal collection), are (optimally) assigned a Persistent Identifier (PID) and hold unique information, e.g. biological and historical, and serve as vouchers for the species they represent. As such, museum collections are valuable time capsules of past and present biodiversity. Their academic and societal value continues to grow, especially in the context of accelerating human-driven environmental change and the unprecedented loss of biodiversity (e.g., Cowie et al. 2022).

Natural history collections have long served as invaluable resources for scientific research in fields such as taxonomy, systematics, and biodiversity studies (e.g., Kahrer and Zettel 2022, Schubert et al. 2023, Enghoff and Akkari 2024, Schwentner and Hethke 2025). In recent years, however, the scope of research utilizing these collections has expanded considerably, moving beyond traditional approaches and leading to a growing interest in scientific collections across a wider range of disciplines, including evolutionary biology, ecology, epidemiology, environmental research or biodiversity assessments (e.g., Pyke and Ehrlich 2010, Card et al. 2021, Hilton et al. 2021, Raxworthy and Smith 2021, Jensen et al. 2022, Fong et al. 2023). This development is largely driven by recent methodological advances that enable the generation of genome-level DNA sequence data (Orlando et al. 2021) and high-resolution 2D and 3D imaging (e.g., Akkari et al. 2015, 2018), as well as the integration of these data with datasets from other databases. At the same time, progress in bioinformatics and computational methods has significantly enhanced the ability to quantitatively analyze such complex datasets and help to increasingly discover and exploit valuable museum collections to answer various scientific questions and bridge with other disciplines (Thomas 1994, Graham et al. 2004, Suarez and Tsutsui 2004). Many of these advanced techniques fall under the umbrella of so-called “-omics” approaches – including, for example, genomics and proteomics.

Genomics has arguably been the primary driver of advances in modern collection-based research, with natural history collections becoming increasingly important sources of genetic material. This, in turn, has significantly heightened awareness of the scientific and societal value of natural history collections. Two major developments promoted this trend. First, natural history museums have established over the past decades cryogenic collections of tissues and DNA, providing valuable sources of high-quality genomic material (Card et al. 2021). Second, the advent of next-generation sequencing (NGS) technologies has opened

up vast historical collections to genomic research by enabling the analysis of low-quality and low quantity DNA (Hofreiter et al. 2015, Orlando et al. 2021). These analyses often included DNA extracted from naturally preserved specimens that are up to hundreds of thousands of years old, i.e., *ancient* DNA (aDNA) or DNA obtained from preserved biological specimens – often housed in museum or herbarium collections – that were not originally collected or stored with molecular analysis in mind, i.e., *historical* DNA (hDNA; definitions adopted from Raxworthy and Smith 2021). Alongside freshly collected, high-quality DNA stored in cryogenic collections, these three categories are referred to as *archival* DNA. Specialized laboratory protocols tailored for short and/or single-stranded DNA fragments – a typical feature of degraded historical DNA, combined with NGS techniques now enable the sequencing of even minute amounts of highly fragmented and degraded DNA, as commonly found in decades- or centuries-old specimens. Additionally, recent advances in bioinformatics allow researchers to detect and correct for DNA degradation and contamination, both of which can confound data analysis and the biological interpretation of genomic information from historical samples (Jónsson et al. 2013, Skoglund et al. 2014, Peltzer et al. 2016, Colella et al. 2020, Davis and Knapp 2024).

In museum collections, type specimens serve as the definitive reference points for species identification and classification, and provide a stable framework for species description. Sequencing historical type specimens enables researchers to directly link modern genetic data – such as DNA barcodes or even whole genomes – to formally described species (e.g., Federhen 2015, Li et al. 2015, Straube et al. 2021, Agne et al. 2022, Sullivan et al. 2022, Letsch et al. 2025) and to improve the quality and coverage of reference databases like the barcode of life database (BOLD; https://boldsystems.org/). Also the methodological expansion and the refinement of computer-assisted techniques fostered a broadening of research fields and questions that could be addressed using museum material, including studies in evolutionary biology (e.g., Palandačić et al. 2023, Strunov et al. 2023), population genetics, archaeology (e.g., Maixner et al. 2021), environmental DNA studies (e.g., Jeunen et al. 2024), ecology (e.g., Newbold 2010), and the assessment of anthropogenic impacts such as climate change and biodiversity loss (e.g., Robin et al. 2022). The combination of datasets from historical and modern sources enabled unique comparisons of temporal patterns in genetic variation and biodiversity. For example, the integration of DNA barcoding data from museum specimens with metabarcoding data from contemporary samples provides increasingly detailed insights into both ecological and evolutionary processes. Such approaches may have practical applications, as shown for biodiversity monitoring (e.g., Pascher et al. 2022) for analysing gut microbiomes (Greiman et al. 2018). Furthermore, natural history collections are increasingly used in medical research contexts, for example in

studies on parasitism and zoonotic diseases and similar research fields (Davis and Knapp 2024, Hämmerle et al. 2024).

Terminology

Emerging research fields at the intersection of traditional museum collections and modern molecular and computational approaches have given rise to a variety of "-omics" neologisms, such as *museum genomics*, *collection genomics, museomics* and *collectomics*, which are increasingly relevant for natural history museums. While these neologisms are often used synonymously, they may refer to distinct yet overlapping concepts. In this article, we provide an overview of the key terms, examine their conceptual interrelations, and present a literature-based survey of their usage. Some terms can essentially be regarded as synonyms, while other ones form a hierarchical system, with broader terms fully encompassing the more specific ones. Following our understanding, we present below the most prominent -omics terms relevant for natural history collections used in our literature search in order of increasing inclusivity.

Museum genomics and collection genomics: These terms specifically refer to the generation and analysis of genomic data – such as whole genomes, mitochondrial genomes (e.g., Dong et al. 2024), DNA barcodes, or single nucleotide polymorphisms (SNPs) – from museum specimens. The terms are typically used when the focus lies exclusively on DNA-based approaches, especially in research fields such as evolutionary biology, systematics, population genetics, or conservation genomics. *Museum genomics* (Rowe et al. 2011) and *collection genomics* (Jin et al. 2020) are often used synonymously. In some instances they have even been used as synonyms for *museomics* (e.g., Card et al. 2021). Yet, the latter term has a broader meaning (see below).

Since archival DNA in museum specimens is typically degraded and fragmented due to the specimens' age or storage conditions (Raxworthy and Smith 2021), specialized laboratory protocols and customized bioinformatic pipelines are required to recover and analyze usable genetic information (Zimmermann et al. 2008, Fulton and Shapiro 2019). Whether the generation of single-gene DNA barcodes qualifies as *museum genomics* is still under debate: barcoding generally targets only a small part of the genome, but when techniques such as genome skimming are used to obtain these sequences, they arguably fall under the broader umbrella of genomic methods. Regardless of methodological scope, the resulting genomic data contribute significantly to the *extended specimen* (Webster 2017, see below; Lendemer et al. 2020), linking molecular, historical, and contextual information and increasing the scientific value of each specimen.

Museomics: This term refers to modern -omics-scale techniques applied to museum specimens for data generation and analysis. While a large share of *museomics* studies focus on genomic data, *museomics* encompasses a wider range of -omics disciplines in addition to genomics. These include proteomics, transcriptomics, metabolomics, and others, as well as specialized bioinformatic tools tailored to deal with the specific challenges of historic and archival molecules from museum specimens. Thus, *museomics* can be considered an umbrella term encompassing methodological approaches often referred to as *museum genomics* or *collection genomics*. Because the focal specimens are often old, fragile, and irreplaceable, standard molecular protocols are frequently adapted to be less or non-destructive, or to overcome specific limitations associated with degraded or limited material (see above). The insights gained through these approaches significantly enrich the data environment of the extended specimen.

According to Lalueza-Fox (2022), *museomics* refers to the "*application of -omics techniques (genomics, paleogenomics, and even paleoproteomics) to previously intractable historical and archival specimens that allow the retrieval of genomic data from extinct or currently declining species*" (Lalueza-Fox 2022, R1214). The term *museomics* was first introduced in 2008 (http://mammoth.psu.edu/museomics.html) in reference to the extinct mammoth Adamy, describing *"the large-scale analysis of the DNA content of museum collections."* It combines "museum" with the suffix "-omics," originally acknowledging its roots in genomics, typically involving NGS followed by genomic analyses. In general, the applied methodologies, which often employ modern NGS techniques, vary widely, ranging from single gene sets (e.g., the DNA barcoding gene cytochrome c oxidase subunit 1 [CO1]), reduced genomic representation approaches, such as ddRAD sequencing, to complete mitochondrial, plastid, or nuclear genomes. However, also non-genetic -omics approaches, based on other organic molecules such as RNA (e.g., Smith et al. 2014) or proteins (e.g., Anderson 2023) are explicitly included in museomics.

Extended specimen: The *extended specimen* as a concept refers to the enhancement of traditional physical specimens through the integration of a wide range of associated data types that collectively increase their scientific value (Webster 2017, Lendemer et al. 2020). These may include additional physical components (e.g., tissue samples, slide preparations), digital data (e.g., high resolution images, x-ray or µCT-scans and DNA sequences), and contextual metadata (e.g., georeferenced collection data, ecological, or functional traits; Lendemer et al. 2020). The extended specimen can also include comprehensive morphological data, which is referred to as morphomics (sensu Son et al. 2023). A crucial requirement is that each specimen is uniquely identifiable via a Persistent

Identifier (PID) to enable reliable linking of the physical object and its digital twin with all associated data (Lendemer et al. 2020), including in data portals such as GBIF (https://www.gbif.org/), GenBank (https://www.ncbi.nlm.nih.gov/genbank/) or BOLD (https://boldsystems.org/).

The idea of combining different types of data around specimens originated in ornithology, where the *extended specimen* was first described as "*a constellation of specimen and data types that jointly add research value to each other*" (Webster 2017, xi). The concept has since been rapidly adopted in the broader natural history community, particularly within CETAF (Consortium of European Taxonomic Facilities; https://cetaf.org/), and plays a central role in the development of the European Research Infrastructure DiSSCo (Distributed System of Scientific Collections: https://www.dissco.eu/; McCormack et al. 2017, Addink et al. 2019, Hardisty et al. 2022). In typical visual representations of the concept (Lendemer et al. 2020, fig. 1), a biological museum object is positioned at the center, while diverse layers of data from various fields are digitally linked to it. These span from genetics (including *museomics*), morphology, geography, and ecology to less commonly integrated aspects such as host-parasite relationships or the historical and political context of the specimen (e.g., collection expeditions, research use, colonial provenance). The extended specimen concept can be adopted for other groups of objects in museum collections from minerals to artefacts, to understand nature and the human-nature-culture relationship.

The terms *cybertype* and *cyberspecimen* have also appeared in literature and are very much in line with the concept of the *extended specimen* as these act as digital counterparts of the physical specimens and type-specimens they correspond to, allowing the corresponding taxa to be included in more research efforts (Faulwetter et al. 2013, Akkari et al. 2015, 2018). Another recent term, *digital extended specimen* (*DES*), appeared in relation to the *extended specimen* and refers to "*the collective representation on the internet of all digital assets referring to a physical specimen*" (Hardisty et al. 2022). The notion of globally digitally linking specimen data from natural science collections to relevant ecological and environmental data from numerous domains is intended to enhance the knowledge on the objects of the scientific collections and transform them into "*digital objects on the internet*" (Hardisty et al. 2022). The term *DES* further stretches the *extended specimen* as a network of digital units. Neither the *extended specimens*, nor their digital replicas are intended to replace physical specimens but add value to them by conferring a superior access to a plethora of data and data network towards creating and managing global digital, virtual scientific collections.

Collectomics: The concept of *collectomics* was first formulated collaboratively by federal research museums of Germany, with the Senckenberg Gesellschaft für Naturforschung playing a leading role in its development. The aim is to extract the maximum scientific and cultural value from collections and link it to other available datasets, such as earth observation, climate and socioeconomic data to address major societal and environmental challenges of the Anthropocene (Sigwart et al. 2025). This includes aligning future collection development with identified knowledge gaps and advancing traditional collection-based research through the incorporation of modern technologies such as molecular genetics, digital imaging, isotope forensics, artificial intelligence and deep learning.

Collectomics refers to the large-scale digital and molecular analysis of natural history collections, emphasizing a holistic, collection-wide perspective rather than focusing on individual specimens or isolated data types. It integrates diverse forms of data – including specimen metadata, digitized images, genomic and other -omic data, as well as associated environmental, ecological, and historical information. At its core, *collectomics* seeks to utilize the collective informational wealth of natural history collections, referred to as the collectome – the totality of all *extended specimens* and their associated data (Misof et al. 2023). Sigwart et al. (2025) defined *collectomics* as *"the multiple ways that voucher specimens can be harnessed for additional datasets, as well as the many ways these data interact synergistically with other research fields including science and the humanities in order to integrate information across scales and generate new knowledge."* Their model builds on the *extended specimen* concept and further integrates data layers such as biodiversity models and environmental or historical data.

Taxonomics: The term *taxonomics* was proposed by Reboleira and Enghoff (2016) as "the discipline of discovering, recognising, describing and classifying biological entities". The authors explain that taxonomics should not only rely on conventional approaches for describing new taxa but also next generation methods, with voucher specimens deposited in scientific collections as a basis. However, this definition largely reflects our general understanding of *integrative taxonomy*. In parallel, "Taxon-OMICs" was used as the name for a priority program (SPP) of the German research foundation (DFG) to foster taxonomic research through the use of omics-scale methods (https://www.taxon-omics.com/). In the scientific literature, the term *taxonomics* is however often used ambiguously and synonymously with *taxonomy* without an explicit reference to omics-scale approaches. Furthermore, it appears mostly in papers with a microbiological focus (e.g., Ettamimi et al. 2019, Blackwell et al. 2020) or in clinical studies, but has not yet been established in the

broader taxonomic literature. We therefore opted not to include this term in the following analyses.

Usage of museum-related -omics terms in literature

To gain an overview of the terminology used in relation to museum collections (see above), we employed a 'snowball research' meta-analysis strategy. We conducted a formalized and restricted literature search based on the keywords: “*museum genomics”*, “*museomics”*, “*extended specimen”*, “*collection genomics”* and “*collectomics”* in the Web of Science (WoS) Core Collection (Clarivate database) on January 14th, 2025, including all publications published up to the end of 2024. Publications were included if the respective keywords appeared in the title, abstract, or keywords section. We stress that the total number of publications dealing with topics related to -omics-scale analyses of museum objects is potentially much larger, but for reproducibility we limited our analysis to those entries that explicitly mention the aforementioned keywords. The associated WoS research areas were downloaded from Clarivate, and both the raw and analyzed data as well as the code for the data analyses and visualizations based on R (R Core Team 2019) and the tidyverse package (Wickham et al. 2019) are archived in the repository: https://github.com/nhmvienna/WoS_Museomics.

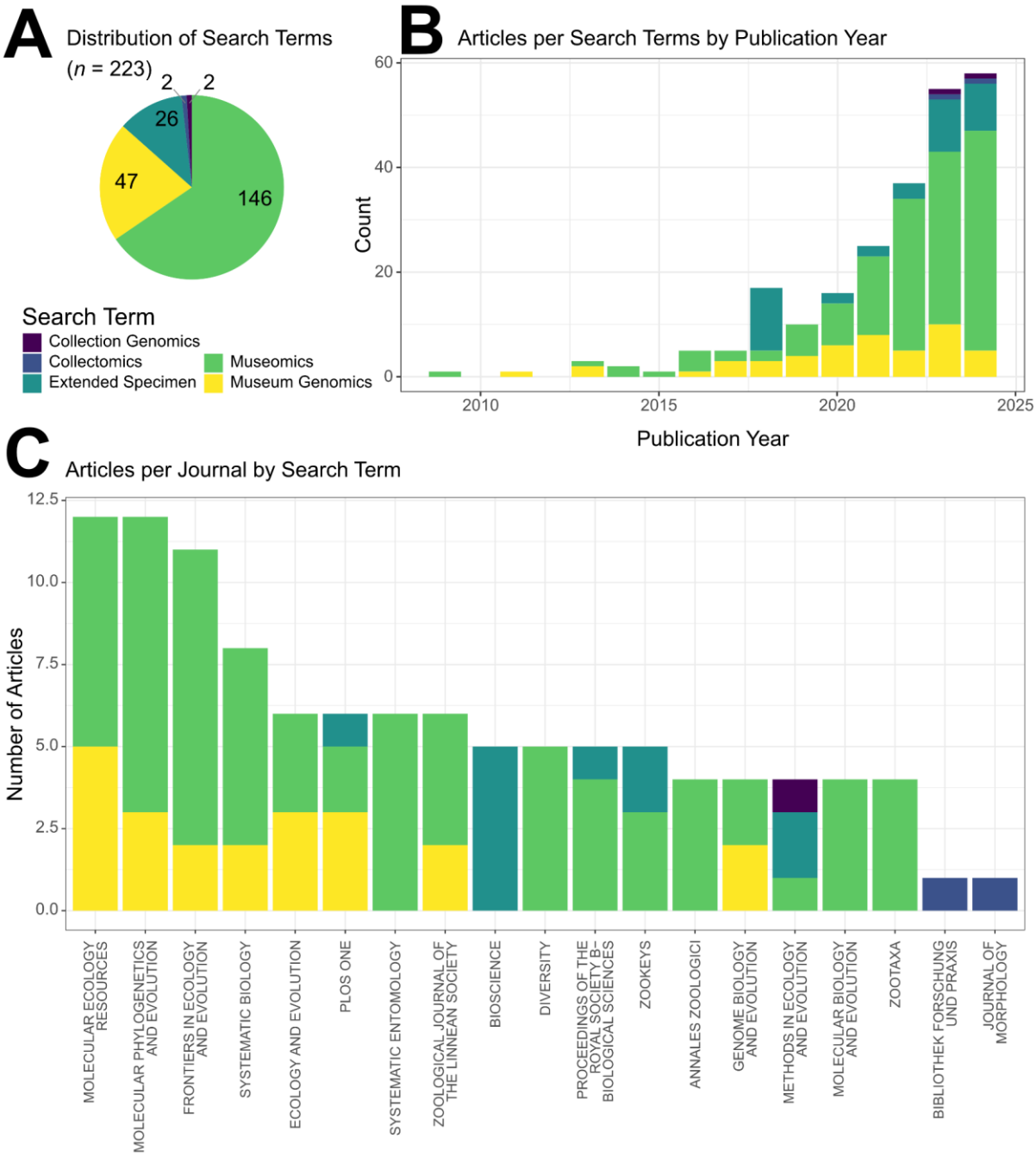

Figure 1: Panel A: Results of a web search based on the Web of Science Core Collection (WoS) based on five search terms. All articles were distributed across 107 journals. The histograms in Panel B and C show article counts that are color-coded per search term and sorted and plotted either by publication year (B) or by scientific journals (C). Journals with less than four publications (n=89) were excluded for the analysis in Panel B.

From the total dataset (n_{Total}= 271), we removed all publications where the terms were used in a clearly different context (mainly *extended specimen* in medical publications), resulting in 223 hits across 107 journals or books (Figure. 1). In our query, the term *museomics* clearly dominated (Figure 1A): *Museomics* (146 hits) was used three times more often than *museum genomics* (47 hits) and *collection genomics* (2 hits) combined. The relatively high

number of articles referring to the *extended specimen* resulted from a special issue *("Extended specimen: emerging frontiers in collections-based ornithological research")* in the book series *"Studies in Avian Biology"* that was published in 2018. Figure 1B shows that the different research fields emerged recently and particularly increased only during the past five years. Especially the usage of the term *museomics* markedly rose after 2018. The highest number of publications (12 each) were published in the journals *Molecular Ecology Resources* and Molecular *Phylogenetics and Evolution*, followed by *Frontiers in Ecology and Evolution* (11), *Systematic Biology* (8) and *Ecology and Evolution* (6) – all high ranking journals with high impact factors (Figure 1C). Following the classification of WoS, most publications dealt with topics in the research fields of evolutionary biology and environmental sciences & ecology, but also in fields specifically focusing on the biology and systematics of organisms such as entomology, zoology, or plant sciences (Figure 2). The comparably high number of articles in the research fields biochemistry and molecular biology indicates ongoing development of new molecular and bioinformatics methods. *Museomics* was the most frequently used term and had clear linkages to biological sciences. *Museum genomics* had a similar profile but was used less frequently. In contrast to *collectomics* and *collection genomics*, articles assigned to *museomics*, *museum genomics* and the *extended specimen* concept were linked to many different research fields (≥15 fields each; Figure 2B), whereas *collectomics* and *collection genomics*, were both limited to only two and five fields, respectively. While WoS linked *collection genomics* mostly to biological research fields, *collectomics* was assigned to information sciences and anatomy/morphology only. However, the small number of nodes linking these two terms within the network is probably due to the low number of publications which precludes a robust assessment to date.

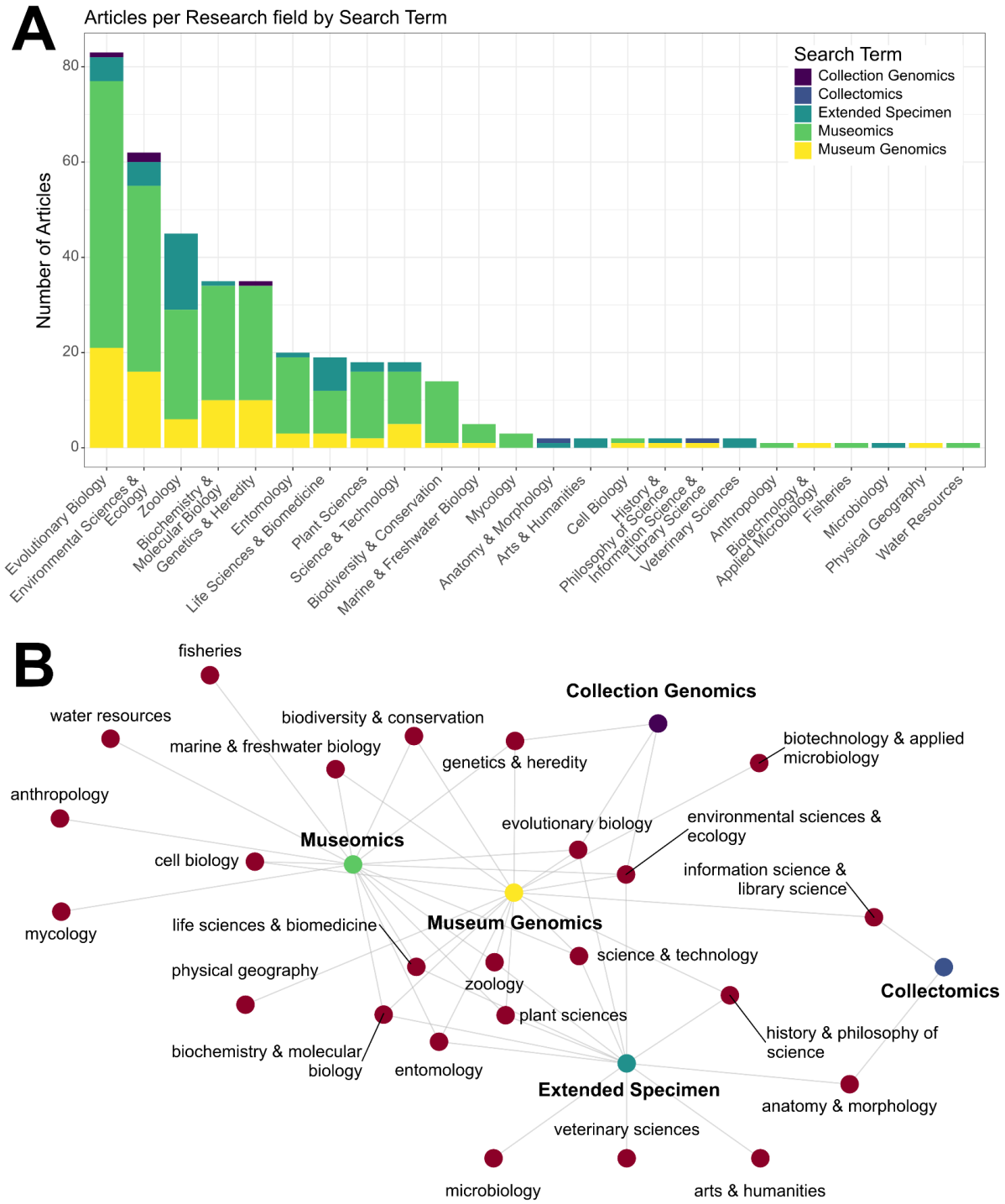

Fig. 2: Panel A: Articles counts per research field (as obtained from WoS) are color-coded according to search term. Note, that some articles were assigned to multiple research fields as also shown in the network in Panel B, which is based on a matrix with shared links between search terms and research fields.

We further manually and qualitatively assessed the preferred analysis methods, genetic marker types and research organism in a subset of all articles that were identified by the keyword *museomics* and found that the vast majority applied shotgun sequencing and target

enrichment for either whole genome analyses of organelles such as mitochondria and plastids or for sequencing ultraconserved elements (UCEs) for phylogenetic inference. Moreover, most studies focused either on insects, plants, or vertebrates such as mammals, birds, amphibians or fishes (Figure 3).

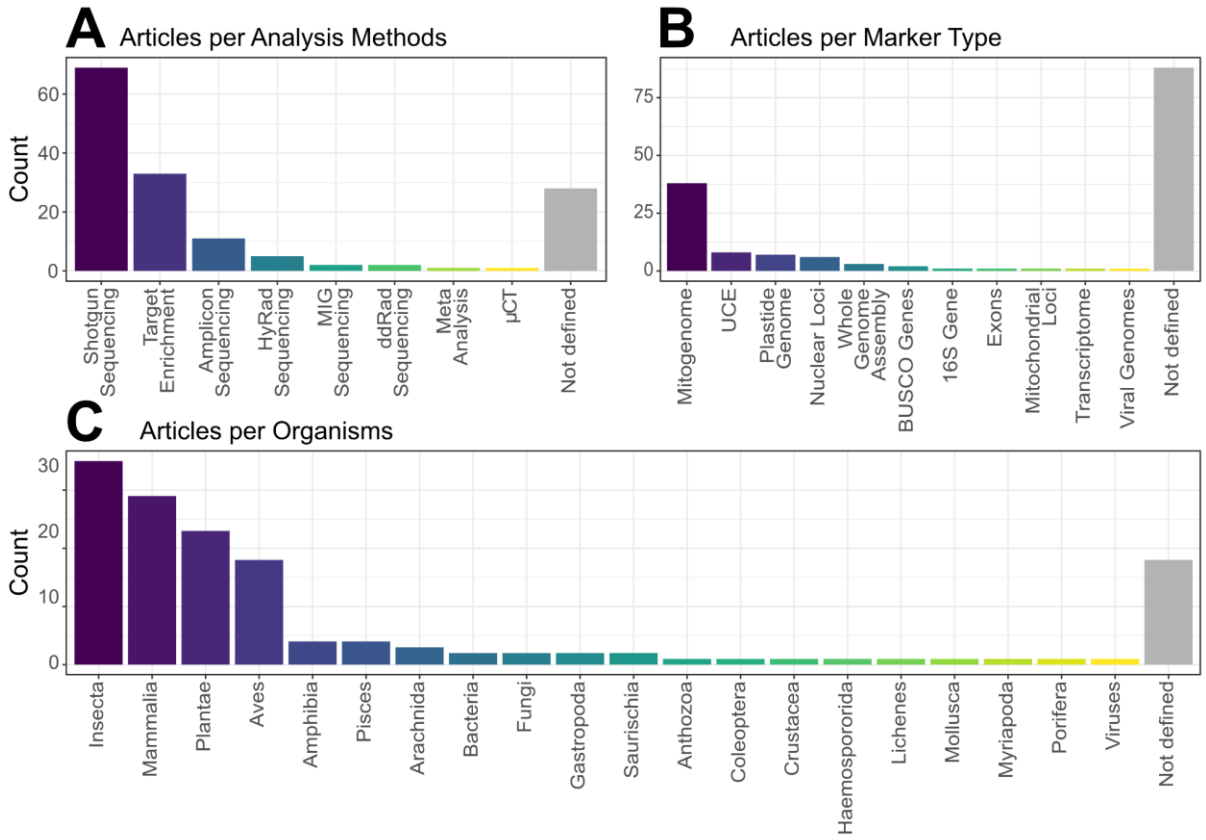

Fig. 3. Histograms showing the numbers of articles describing specific analysis methods (Panel A), investigated genetic marker types (Panel B) and organismal groups (at different hierarchical levels; Panel C). The dataset is based on a hand-curated qualitative analysis of all 146 articles that were identified with the search term "museomics". In each panel, we indicated articles which we were unable to assign to one of the categories in grey. UCE - Ultraconserved Elements; MIG- sequencing - multiplexed ISSR genotyping by sequencing (Suyama and Matsuki 2015)*; ddRAD - double digest restriction-site associated DNA; hyRAD - hybridization restriction-site associated DNA*

Together, these analyses underpin the recent dawn and growth of -omics-scale analyses of museum material and highlight the diversity in research fields, applications and focal species investigated with these new approaches. Moreover, the wealth of methods-related publications indicates that these fields are still under very active development and will be further boosted in the future by ongoing innovation through methodological advances and novel interdisciplinary research questions.

Which term for which purpose?

Our literature review has shown that terms that have a strong methodological focus, such as *museomics*, *museum genomics*, *collection genetics,* and *collection genomics*, are often used synonymously (for example, see Card et al. 2021). *Museomics* is most commonly used and its usage as a keyword in scientific literature has particularly increased over the last years. Conceptually it is the most inclusive among these terms as it is not strictly limited to genetics or genomics but also encompasses other -omic techniques that investigate biomolecules. To avoid inflationary usage of a plethora of largely synonymous terms or the creation of new neologisms (e.g., Eisen 2012), we strongly recommend to use *museomics* as the technical term to describe all research that investigates genetic data or other biomolecules from museum specimens of any age, i.e., containing archival DNA (Raxworthy and Smith 2021).

Collectomics has not yet gained widespread acceptance as a standard term within the international scientific community. However, the underlying concepts, ideas and visions are crucial for the integration of collection-based objects and their associated data with other large-scale, multidimensional data sets like climate, land-use, historical, or ethnographic data to approach and tackle global challenges (Sigwart et al. 2025). Herein the *extended specimen* concept takes a central role as it links physical objects with derived and associated data (cf. Hardisty et al. 2022). To harness the wealth of information stored in museum collections it is essential that the data is digitized and made publicly available following the FAIR (Wilkinson et al. 2016) and CARE (https://www.gida-global.org/care) principles. As of now, only a small fraction of natural history collections has been properly digitized (Sigwart et al. 2025) and highly enriched data (e.g. genetic or tomographic data) are available for only a tiny subset. Generating, providing, and curating the data is a huge and complex task for all natural history collections that requires further investments in the infrastructure as well as specialized personnel.

One aspect that is often neglected in these discussions is the acquisition of new material into museum collections. Of course, cryogenic collections can help to store genomic resources for future generations in the form of yet unexplored archival DNA (Raxworthy and Smith 2021). But natural history collections also serve as repositories of natural history objects that are newly collected and utilized in scientific research. Keeping these specimens as vouchers in the natural history collections, on one hand, ensures the reproducibility of the research (Davis and Knapp 2024), for example, in the form of DNA barcode reference vouchers, whose species identification can be verified by other taxonomic experts. On the other hand, these vouchers represent extended specimens that are already enriched with additional data and thus directly contribute to the goals of *collectomics*.

In summary, we therefore recommend to focus on three main terms that describe -omics scale analyses of museum objects at three different hierarchical levels: (1) *museomics* includes analyses methods that focus on all types of biomolecules in museum specimens of various age and degradation level, whereas the (2) *extended specimen* concept describes approaches to integrate a physical object with all kinds of digital metadata that allow for holistic and interdisciplinary analyses. Finally, (3) *collectomics* intersects databases that build upon extended specimens with quantitative datasets from different domains which ultimately facilitates analyses at different hierarchical levels beyond the individual object.

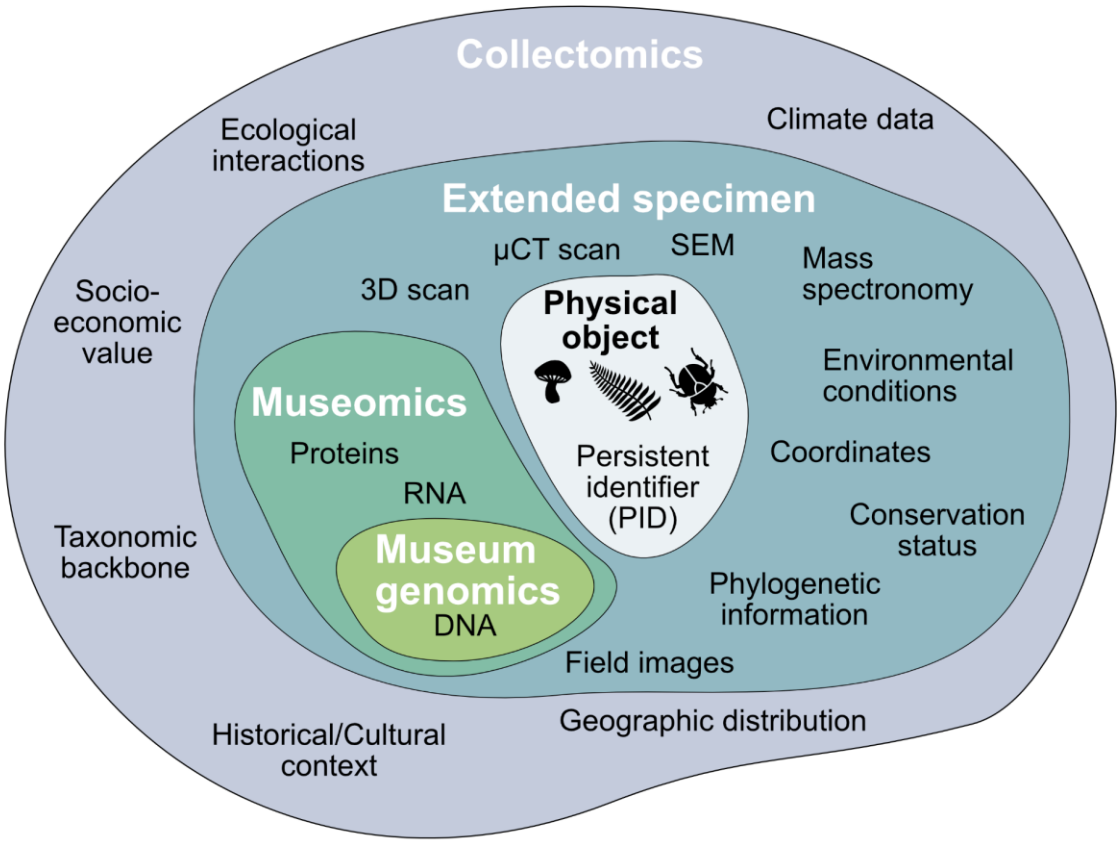

Figure 4. Diagram showing hierarchical data layers and corresponding -omics-scale approaches, discussed in this manuscript, to analyse and link datasets to physical museum objects.

Perspectives from a natural history museum – linking data and specimens

The curatorial work at the museum plays an essential role in building an extended specimen model, which does not only refer to the creation of a digital twin, but also includes the assignment of a permanent digital identifier (PID) and additional metadata such as

coordinates and collection date as well as additional morphological and genetic information. These further include, for example, illustrations of key taxonomic characters, volumetric data or DNA barcodes. It is thus pivotal to commit to a museum-wide strategy for building extended specimen datasets using standardized data formats following Open Science principles to pave the ground for omics-scale analyses at different levels as presented in this document. This facilitates collaborative work within and among research institutions that can fully benefit from the recent technological advances and will only expand the relevance and impact of the natural history collections as main resources for scientific research, and education. An idea of a “Global Museum” to science and society, emerging from linking digital natural history resources to answer global biological, environmental and societal questions has already been proposed (Bakker et al. 2020).

At the Natural History Museum Vienna (NHMW), we are committed to data analysis of museum specimens at all aforementioned -omics-scale levels, to connect the data to the physical objects (following the *extended specimen* concept), and to make these data publicly available following Open Science principles (Vohland et al. 2022). For example, we use *museomics* approaches ranging from DNA barcoding (e.g., Bronstein et al. 2017, 2018) to shotgun sequencing (Palandačić et al. 2023) of historic DNA and quantitative bioinformatics data analysis to address questions at the interface of taxonomic classification (Bronstein et al. 2017), biodiversity assessment and evolutionary biology in diverse organisms and their endosymbionts (Strunov et al. 2023). Large-scale DNA barcoding, for example, is pioneered and heavily applied by the Austrian Barcode of Life project (ABOL; https://www.abol.ac.at/), which is coordinated by the NHMW and aims at building a comprehensive database of DNA barcodes of Austrian species and storing genetic data in respective repositories such as GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and BOLD. DNA barcoding has become a standard tool for linking specimens identified by taxonomic experts to their genetic information. The requirement for reproducibility through voucher specimens deposited in collections was emphasized from the outset (Hebert et al. 2004). The open-access DNA database BOLD (Ratnasingham et al. 2024) facilitates species identification by non-specialists and enables non-invasive detection of species via environmental DNA (eDNA). Beyond these applications, a DNA barcode sequence represents a key genetic attribute of the *extended specimen*. Particularly valuable are barcodes derived from type material. As such, BOLD plays a central role in the transformation of biological collections into interdisciplinary research infrastructures (Koureas et al. 2024).

At the national level, the OSCA (Open Scientific Collections Austria) consortium was established with the goal of digitizing the natural historic collections housed in Austrian

museums, universities and other institutions, such as monasteries (https://osca.science/). The NHMW is setting up OSCA as the national node for DiSSCo (Distributed System of Scientific Collections; https://www.dissco.eu/). To bring together and link all data of an extended specimen, we are currently building an institutional collection database, DIVINA (Digital Information for VIenna NAtural history museum; https://divina.nhm.at/home), which is based on the Darwin core data standard (Wieczorek et al. 2012) and allows us to integrate and link digital copies of collection items and digital metadata catalogues from the various collections at the NHMW. The NHMW has also set up an online data repository (https://datarepository.nhm.at/) to store research data, which are linked to the respective physical object in DIVINA, and to make this data publicly available. Moreover, following the FAIR principles, DIVINA provides an OAI-PMH API (https://oai.nhm.at/) which allows integrating the datasets into comprehensive national (such as Kulturpool; https://kulturpool.at/) and international portals such as Europeana, GBIF or EOSC. These linked data facilitate *collectomics* analyses that intersect *extended specimen* datasets with external data from earth observation, climatic, or socioeconomic databases. As part of the Horizon-Europe funded FAIRiCUBE (https://fairicube.nilu.no/) project, we are, for example, contributing to a digital hub which allows us to carry out integration of quantitative datasets from different domains. Using these new tools, we conduct case studies that link comprehensive earth observation with biodiversity data to investigate the anthropogenic impact on biodiversity and species distribution.

Authors Contributions

MK: Conceptualization, Data Curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing; MS: Conceptualization, Data Curation, Validation, Visualization, Writing - original draft, Writing - review & editing, EH: Conceptualization, Data Curation, Methodology, Validation, Writing - original draft, Writing - review & editing; NA: Conceptualization, Methodology, Validation, Writing - review & editing; AK: Conceptualization, Methodology, Validation, Writing - review & editing; LK: Conceptualization, Methodology, Validation, Writing - review & editing; AP: Conceptualization, Methodology, Validation, Writing - review & editing; KV: Conceptualization, Data Curation, Formal analysis, Methodology, Project administration, Validation, Visualization, Writing - original draft, Writing - review & editing;

Data availability

All data and code that underlie the analysis and figures presented in this manuscript can be found at https://github.com/nhmvienna/WoS_Museomics. Github CoPilot (GPT-4o) was used for code completion and code auditing of the *R* code written in this project. All URLs in this manuscript have been last visited on 04/06/2025.

Acknowledgments

We are very grateful to all our colleagues at the Natural History Museum of Vienna for the ongoing and fruitful discussions on the topics presented here.

References

Addink W, Koureas D, Rubio A (2019) DiSSCo as a New Regional Model for Scientific Collections in Europe. Biodiversity Information Science and Standards 3: e37502. https://doi.org/10.3897/biss.3.37502

Agne S, Preick M, Straube N, Hofreiter M (2022) Simultaneous Barcode Sequencing of Diverse Museum Collection Specimens Using a Mixed RNA Bait Set. Frontiers in Ecology and Evolution 10. https://doi.org/10.3389/fevo.2022.909846

Akkari N, Enghoff H, Metscher BD (2015) A New Dimension in Documenting New Species: High-Detail Imaging for Myriapod Taxonomy and First 3D Cybertype of a New Millipede Species (Diplopoda, Julida, Julidae). PLOS ONE 10: e0135243. https://doi.org/10.1371/journal.pone.0135243

Akkari N, Ganske A-S, Komerički A, Metscher B (2018) New avatars for Myriapods: Complete 3D morphology of type specimens transcends conventional species description (Myriapoda, Chilopoda). PLOS ONE 13: e0200158. https://doi.org/10.1371/journal.pone.0200158

Anderson LA (2023) A chemical framework for the preservation of fossil vertebrate cells and soft tissues. Earth-Science Reviews 240: 104367. https://doi.org/10.1016/j.earscirev.2023.104367

Bakker FT, Antonelli A, Clarke JA, Cook JA, Edwards SV, Ericson PGP, Faurby S, Ferrand N, Gelang M, Gillespie RG, Irestedt M, Lundin K, Larsson E, Matos-Maraví P, Müller J, Proschwitz T von, Roderick GK, Schliep A, Wahlberg N, Wiedenhoeft J, Källersjö M (2020) The Global Museum: natural history collections and the future of evolutionary science and public education. PeerJ 8: e8225. https://doi.org/10.7717/peerj.8225

Blackwell JM, Fakiola M, Singh OP (2020) Genetics, Transcriptomics and Meta-Taxonomics in Visceral Leishmaniasis. Frontiers in Cellular and Infection Microbiology 10. https://doi.org/10.3389/fcimb.2020.590888

Bronstein O, Georgopoulou E, Kroh A (2017) On the distribution of the invasive long-spined echinoid Diadema setosum and its expansion in the Mediterranean Sea. Marine Ecology Progress Series 583: 163–178. https://doi.org/10.3354/meps12348

Bronstein O, Kroh A, Haring E (2018) Mind the gap! The mitochondrial control region and its power as a phylogenetic marker in echinoids. BMC Evolutionary Biology 18: 80. https://doi.org/10.1186/s12862-018-1198-x

Card DC, Shapiro B, Giribet G, Moritz C, Edwards SV (2021) Museum Genomics. Annual Review of Genetics 55: 633–659. https://doi.org/10.1146/annurev-genet-071719-020506

Colella JP, Tigano A, MacManes MD (2020) A linked-read approach to museomics: Higher quality de novo genome assemblies from degraded tissues. Molecular Ecology Resources 20: 856–870. https://doi.org/10.1111/1755-0998.13155

Cowie RH, Bouchet P, Fontaine B (2022) The Sixth Mass Extinction: fact, fiction or speculation? Biological Reviews 97: 640–663. https://doi.org/10.1111/brv.12816

Davis CC, Knapp S (2024) Exploring biodiversity through museomics. Nature Reviews Genetics. https://doi.org/10.1038/s41576-024-00801-2

Dong J, Liu Y, Tan MK, Wahab RA, Nattier R, Chifflet-Belle P, Robillard T (2024) Museomics allows comparative analyses of mitochondrial genomes in the family Gryllidae (Insecta, Orthoptera) and confirms its phylogenetic relationships. PeerJ 12: e17734. https://doi.org/10.7717/peerj.17734

Eisen JA (2012) Badomics words and the power and peril of the ome-meme. GigaScience 1: 6. https://doi.org/10.1186/2047-217X-1-6

Enghoff H, Akkari N (2024) A new species of Lophostreptus Cook, 1895 discovered among syntypes of L. regularis Attems, 1909 (Diplopoda, Spirostreptida, Spirostreptidae). ZooKeys 1188: 265–274. https://doi.org/10.3897/zookeys.1188.115802

Ettamimi S, Carlier JD, Cox CJ, Elamine Y, Hammani K, Ghazal H, Costa MC (2019) A meta-taxonomic investigation of the prokaryotic diversity of water bodies impacted by acid mine drainage from the São Domingos mine in southern Portugal. Extremophiles 23: 821–834. https://doi.org/10.1007/s00792-019-01136-1

Faulwetter S, Vasileiadou A, Kouratoras M, Thanos Dailianis null, Arvanitidis C (2013) Micro-computed tomography: Introducing new dimensions to taxonomy. ZooKeys: 1–45. https://doi.org/10.3897/zookeys.263.4261

Federhen S (2015) Type material in the NCBI Taxonomy Database. Nucleic Acids Research 43: D1086–D1098. https://doi.org/10.1093/nar/gku1127

Fong JJ, Blom MPK, Aowphol A, McGuire JA, Sutcharit C, Soltis PS (2023) Editorial: Recent advances in museomics: revolutionizing biodiversity research. Frontiers in Ecology and Evolution 11. https://doi.org/10.3389/fevo.2023.1188172

Fulton TL, Shapiro B (2019) Setting Up an Ancient DNA Laboratory. In: Shapiro B, Barlow A, Heintzman PD, Hofreiter M, Paijmans JLA, Soares AER (Eds), Ancient DNA: Methods and Protocols. Springer, New York, NY, 1–13. https://doi.org/10.1007/978-1-4939-9176-1_1

Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology & Evolution 19: 497–503. https://doi.org/10.1016/j.tree.2004.07.006

Greiman SE, Cook JA, Tkach VV, Hoberg EP, Menning DM, Hope AG, Sonsthagen SA, Talbot SL (2018) Museum metabarcoding: A novel method revealing gut helminth communities of small mammals across space and time. International Journal for Parasitology 48: 1061–1070. https://doi.org/10.1016/j.ijpara.2018.08.001

Hämmerle M, Guellil M, Trgovec-Greif L, Cheronet O, Sawyer S, Ruiz-Gartzia I, Lizano E, Rymbekova A, Gelabert P, Bernardi P, Han S, Rattei T, Schuenemann VJ, Marques-Bonet T, Guschanski K, Calvignac-Spencer S, Pinhasi R, Kuhlwilm M (2024) Screening great ape museum specimens for DNA viruses. Scientific Reports 14: 29806. https://doi.org/10.1038/s41598-024-80780-w

Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, Rabeler RK, Bates J, Bentley A, Fortes JAB, Hansen S, Macklin JA, Mast AR, Miller JT, Monfils AK, Paul DL, Wallis E, Webster M (2022) Digital Extended Specimens: Enabling an Extensible Network of Biodiversity Data Records as Integrated Digital Objects on the Internet. BioScience 72: 978–987. https://doi.org/10.1093/biosci/biac060

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of Birds through DNA Barcodes. PLOS Biology 2: e312. https://doi.org/10.1371/journal.pbio.0020312

Hilton EJ, Watkins-Colwell GJ, Huber SK (2021) The Expanding Role of Natural History Collections. Ichthyology & Herpetology 109: 379–391. https://doi.org/10.1643/t2020018

Hofreiter M, Paijmans JLA, Goodchild H, Speller CF, Barlow A, Fortes GG, Thomas JA, Ludwig A, Collins MJ (2015) The future of ancient DNA: Technical advances and conceptual shifts. BioEssays 37: 284–293. https://doi.org/10.1002/bies.201400160

Jensen EL, Díez-del-Molino D, Gilbert MTP, Bertola LD, Borges F, Cubric-Curik V, de Navascués M, Frandsen P, Heuertz M, Hvilsom C, Jiménez-Mena B, Miettinen A, Moest M, Pečnerová P, Barnes I, Vernesi C (2022) Ancient and historical DNA in conservation policy. Trends in Ecology & Evolution 37: 420–429. https://doi.org/10.1016/j.tree.2021.12.010

Jeunen G-J, Mills S, Lamare M, Duffy GA, Knapp M, Stanton J-AL, Mariani S, Treece J, Ferreira S, Durán-Vinet B, Zavodna M, Gemmell NJ (2024) Unlocking Antarctic molecular time-capsules – Recovering historical environmental DNA from museum-preserved sponges. Molecular Ecology Resources 24: e14001. https://doi.org/10.1111/1755-0998.14001

Jin M, Zwick A, Ślipiński A, de Keyzer R, Pang H (2020) Museomics reveals extensive cryptic diversity of Australian prionine longhorn beetles with implications for their classification and conservation. Systematic Entomology 45: 745–770. https://doi.org/10.1111/syen.12424

Johnson KR, Owens IFP, THE GLOBAL COLLECTION GROUP (2023) A global approach for natural history museum collections. Science 379: 1192–1194. https://doi.org/10.1126/science.adf6434

584 Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L (2013) mapDamage2.0: fast
585 approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics
586 29: 1682–1684. https://doi.org/10.1093/bioinformatics/btt193

587 Kahrer A, Zettel H (2022) Type material of *Leucaspis pusilla* LÖW, 1883 (Insecta:
588 Hemiptera: Diaspididae) recovered in the Natural History Museum Vienna. Annalen
589 des Naturhistorischen Museums in Wien, Serie B 124: 187–192.

590 Koureas D, Livermore L, Addink W, Alonso E, Alonso J, Casino A, Dusoulier F, Ferreira V,
591 Grieb J, Groom Q, Islam S, Kõljalg U, Lymer G, Marhold K, Paleco C, Pijls S, Scory
592 S, Scott B, Weiland C, Worley K (2024) DiSSCo Transition Abridged Grant Proposal.
593 Research Ideas and Outcomes 10: e118241. https://doi.org/10.3897/rio.10.e118241

594 Lalueza-Fox C (2022) Museomics. Current Biology 32: R1214–R1215.
595 https://doi.org/10.1016/j.cub.2022.09.019

596 Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, LeVan K, Bates J,
597 Jennings D, Contreras D, Lagomarsino L, Mabee P, Ford LS, Guralnick R, Gropp
598 RE, Revelez M, Cobb N, Seltmann K, Aime MC (2020) The Extended Specimen
599 Network: A Strategy to Enhance US Biodiversity Collections, Promote Research and
600 Education. BioScience 70: 23–30. https://doi.org/10.1093/biosci/biz140

601 Letsch H, Greve C, Hundsdoerfer AK, Irisarri I, Moore JM, Espeland M, Wanke S, Arifin U,
602 Blom MPK, Corrales C, Donath A, Fritz U, Köhler G, Kück P, Lemer S, Mengual X,
603 Salas NM, Meusemann K, Palandačić A, Printzen C, Sigwart JD, Silva-Brandão KL,
604 Simões M, Stange M, Suh A, Szucsich N, Tilic E, Töpfer T, Böhne A, Janke A, Pauls
605 S (2025) Type genomics: a Framework for integrating Genomic Data into Biodiversity
606 and Taxonomic research. Systematic Biology: syaf040.
607 https://doi.org/10.1093/sysbio/syaf040

608 Li C, Corrigan S, Yang L, Straube N, Harris M, Hofreiter M, White WT, Naylor GJP (2015)
609 DNA capture reveals transoceanic gene flow in endangered river sharks.
610 Proceedings of the National Academy of Sciences 112: 13302–13307.
611 https://doi.org/10.1073/pnas.1508735112

612 Maixner F, Sarhan MS, Huang KD, Tett A, Schoenafinger A, Zingale S, Blanco-Míguez A,
613 Manghi P, Cemper-Kiesslich J, Rosendahl W, Kusebauch U, Morrone SR,
614 Hoopmann MR, Rota-Stabelli O, Rattei T, Moritz RL, Oeggl K, Segata N, Zink A,
615 Reschreiter H, Kowarik K (2021) Hallstatt miners consumed blue cheese and beer
616 during the Iron Age and retained a non-Westernized gut microbiome until the
617 Baroque period. Current Biology 31: 5149-5162.e6.
618 https://doi.org/10.1016/j.cub.2021.09.031

619 McCormack JE, Rodríguez-Gómez F, Tsai WLE, Faircloth BC (2017) Transforming Museum
620 Specimens into Genomic Resources. In: The Extended Specimen. CRC Press.

621 Misof B, Tockner K, Trischler H (2023) Die Bibliothek des biologischen, technischen und
622 kulturellen Wissens – Warum brauchen wir eine integrierte Sammlungsinfrastruktur?
623 Bibliothek Forschung und Praxis 47: 495–502. https://doi.org/10.1515/bfp-2023-0054

624 Newbold T (2010) Applications and limitations of museum data for conservation and
625 ecology, with particular attention to species distribution models. Progress in Physical
626 Geography: Earth and Environment 34: 3–22.
627 https://doi.org/10.1177/0309133309355630

Orlando L, Allaby R, Skoglund P, Der Sarkissian C, Stockhammer PW, Ávila-Arcos MC, Fu Q, Krause J, Willerslev E, Stone AC, Warinner C (2021) Ancient DNA analysis. Nature Reviews Methods Primers 1: 1–26. https://doi.org/10.1038/s43586-020-00011-0

Palandačić A, Kapun M, Greve C, Schell T, Kirchner S, Kruckenhauser L, Szucsich N, Bogutskaya N (2023) From historical expedition diaries to whole genome sequencing: A case study of the likely extinct Red Sea torpedo ray. Zoologica Scripta.

Pascher K, Švara V, Jungmeier M (2022) Environmental DNA-Based Methods in Biodiversity Monitoring of Protected Areas: Application Range, Limitations, and Needs. Diversity 14: 463. https://doi.org/10.3390/d14060463

Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, Nieselt K (2016) EAGER: efficient ancient genome reconstruction. Genome Biology 17: 60. https://doi.org/10.1186/s13059-016-0918-z

Pyke GH, Ehrlich PR (2010) Biological collections and ecological/environmental research: a review, some observations and a look to the future. Biological Reviews 85: 247–266. https://doi.org/10.1111/j.1469-185X.2009.00098.x

R Core Team (2019) R Foundation for Statistical Computing. Available from: https://www.R-project.org/.

Ratnasingham S, Wei C, Chan D, Agda J, Agda J, Ballesteros-Mejia L, Boutou HA, El Bastami ZM, Ma E, Manjunath R, Rea D, Ho C, Telfer A, McKeowan J, Rahulan M, Steinke C, Dorsheimer J, Milton M, Hebert PDN (2024) BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data. Methods in Molecular Biology (Clifton, N.J.) 2744: 403–441. https://doi.org/10.1007/978-1-0716-3581-0_26

Raxworthy CJ, Smith BT (2021) Mining museums for historical DNA: advances and challenges in museomics. Trends in Ecology & Evolution 36: 1049–1060. https://doi.org/10.1016/j.tree.2021.07.009

Robin M, Ferrari G, Akgül G, Münger X, von Seth J, Schuenemann VJ, Dalén L, Grossen C (2022) Ancient mitochondrial and modern whole genomes unravel massive genetic diversity loss during near extinction of Alpine ibex. Molecular Ecology 31: 3548–3565. https://doi.org/10.1111/mec.16503

Rowe KC, Singhal S, Macmanes MD, Ayroles JF, Morelli TL, Rubidge EM, Bi K, Moritz CC (2011) Museum genomics: low-cost and high-accuracy genetic data from historical specimens. Molecular Ecology Resources 11: 1082–1092. https://doi.org/10.1111/j.1755-0998.2011.03052.x

Schubert H, Duda M, Eschner A, Weigand E, Kruckenhauser L (2023) DNA barcoding as a tool to monitor the diversity of endangered spring snails in an Austrian National Park. Biodiversity Data Journal 11: e91496. https://doi.org/10.3897/BDJ.11.e91496

Schwentner M, Hethke M (2025) Revision of the Australian Ozestheria Schwentner & Richter, 2015 (Crustacea: Branchiopoda: Spinicaudata) fauna, with the descriptions of 27 new species. European Journal of Taxonomy 992: 1–172. https://doi.org/10.5852/ejt.2025.992.2905

Sigwart JD, Schleuning M, Brandt A, Pfenninger M, Saeedi H, Borsch T, Häffner E, Lücking R, Güntsch A, Trischler H, Töpfer T, Wesche K, Consortium C (2025) Collectomics – towards a new framework to integrate museum collections to address global challenges. Natural History Collections and Museomics 2: 1–20. https://doi.org/10.3897/nhcm.2.148855

Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J, Jakobsson M (2014) Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proceedings of the National Academy of Sciences 111: 2229–2234. https://doi.org/10.1073/pnas.1318934111

Smith O, Clapham A, Rose P, Liu Y, Wang J, Allaby RG (2014) A complete ancient RNA genome: identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus. Scientific Reports 4: 4003. https://doi.org/10.1038/srep04003

Son R, Yamazawa K, Oguchi A, Suga M, Tamura M, Yanagita M, Murakawa Y, Kume S (2023) Morphomics via next-generation electron microscopy. Journal of Molecular Cell Biology 15: mjad081. https://doi.org/10.1093/jmcb/mjad081

Straube N, Lyra ML, Paijmans JLA, Preick M, Basler N, Penner J, Rödel M-O, Westbury MV, Haddad CFB, Barlow A, Hofreiter M (2021) Successful application of ancient DNA extraction and library construction protocols to museum wet collection specimens. Molecular Ecology Resources 21: 2299–2315. https://doi.org/10.1111/1755-0998.13433

Strunov A, Kirchner S, Schindelar J, Kruckenhauser L, Haring E, Kapun M (2023) Historic museum samples provide evidence for a recent replacement of Wolbachia types in European Drosophila melanogaster. Molecular Biology and Evolution: msad258. https://doi.org/10.1093/molbev/msad258

Suarez AV, Tsutsui ND (2004) The Value of Museum Collections for Research and Society. BioScience 54: 66–74. https://doi.org/10.1641/0006-3568(2004)054[0066:TVOMCF]2.0.CO;2

Sullivan JP, Hopkins CD, Pirro S, Peterson R, Chakona A, Mutizwa TI, Mulelenu CM, Alqahtani FH, Vreven E, Dillman CB (2022) Mitogenome recovered from a 19 th Century holotype by shotgun sequencing supplies a generic name for an orphaned clade of African weakly electric fishes (Osteoglossomorpha, Mormyridae). ZooKeys 1129: 163–196. https://doi.org/10.3897/zookeys.1129.90287

Suyama Y, Matsuki Y (2015) MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. Scientific Reports 5: 16963. https://doi.org/10.1038/srep16963

Thomas RH (1994) Analysis of DNA from natural history museum collections. In: Schierwater B, Streit B, Wagner GP, DeSalle R (Eds), Molecular Ecology and Evolution: Approaches and Applications. Birkhäuser, Basel, 311–321. https://doi.org/10.1007/978-3-0348-7527-1_19

Vohland K, Eichert S, Fiedler S, Kapun M, Kroh A, Mehu-Blantar I, Ott I, Rainer H, Schwentner M, Zimmermann E (2022) Open Science in Museums -Strategy of the Naturhistorisches Museum Wien (NHMW): The benefits of openness. Zenodo https://doi.org/10.5281/zenodo.6505274

Webster MS (Ed.) (2017) The Extended Specimen: Emerging Frontiers in Collections-Based Ornithological Research. CRC Press, Boca Raton, 252 pp. https://doi.org/10.1201/9781315120454

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019) Welcome to the Tidyverse. Journal of Open Source Software 4: 1686. https://doi.org/10.21105/joss.01686

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLOS ONE 7: e29715. https://doi.org/10.1371/journal.pone.0029715

Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. https://doi.org/10.1038/sdata.2016.18

Zimmermann J, Hajibabaei M, Blackburn DC, Hanken J, Cantin E, Posfai J, Evans TC (2008) DNA damage in preserved specimens and tissue samples: a molecular assessment. Frontiers in Zoology 5: 18. https://doi.org/10.1186/1742-9994-5-18