# The Global Museum: natural history collections and the future of evolutionary biology and public education

**Freek T. Bakker** Corresp., 1 , **Alexandre Antonelli** 2 , **Julia Clarke** 3 , **Joseph A Cook** 4 , **Scott V Edwards** 5 , **Per GP Ericson** 6 , **Søren Faurby** 7 , **Nuno Ferrand** 8 , **Magnus Gelang** 9 , **Rosemary G Gillespie** 10 , **Martin Irestedt** 6 , **Kennet Lundin** 9 , **Ellen Larsson** 7 , **Pável Matos-Maraví** 11 , **Johannes Müller** 12 , **Ted von Proschwitz** 9 , **George K Roderick** 10 , **Alexander Schliep** 13 , **Niklas Wahlberg** 14 , **John Wiedenhoeft** 13 , **Mari Källersjö** 15

1 Biosystematics Group, Wageningen University and Research, Wageningen, Netherlands
2 Department of Science, Royal Botanic Gardens, Kew, Richmond, United Kingdom
3 Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, United States
4 Museum of Southwestern Biology, Department of Biology, University of New Mexico, Albuquerque, New Mexico, United States
5 Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts, United States
6 Dept. of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden
7 Department of Biological and Environmental Sciences, Gothenburg Global Biodiversity Centre, University of Gothenburg, Gothenburg, Sweden
8 Museu de História Natural e da Ciência, Universidade do Porto, Porto, Portugal
9 Department of Zoology, Gothenburg Natural History Museum, Gothenburg, Sweden
10 Essig Museum of Entomology, Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, California, United States
11 Department of Ecology, Institute of Entomology ASCR, České Budějovice, Czechia
12 Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Museum für Naturkunde, Berlin, Germany
13 Department of Computer Science and Engineering, University of Gothenburg, Gothenburg, Sweden
14 Department of Biology, Lund University, Lund, Sweden
15 Gothenburg Botanical Garden, Göteborg, Sweden

Corresponding Author: Freek T. Bakker
Email address: freek.bakker@wur.nl

Natural history museums are unique spaces for interdisciplinary research and for educational innovation. Through extensive exhibits and public programming and by hosting rich communities of amateurs, students, and researchers at all stages of their careers, they provide a place-based window to focus on integration of science and discovery, as well as a locus for community engagement. At the same time, like a synthesis radio telescope, when joined together through emerging digital resources, the global community of museums (the ‘Global Museum’) is more than the sum of its parts, allowing insights and answers to diverse biological, environmental, and societal questions at the global scale, across eons of time, and spanning vast diversity across the Tree of Life. We argue that, whereas natural history collections and museums began with a focus on describing the diversity and peculiarities of species on Earth, they are now increasingly leveraged in new ways that significantly expand their impact and relevance. These new directions include the possibility to ask new, often interdisciplinary questions in basic and

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

applied science; inform biomimetic design; and even provide solutions to climate change, global health and food security challenges. As institutions, they are incubators for cutting-edge research in biology and simultaneously protect core infrastructure for present and future societal needs. In this perspective, we discuss challenges to the realization of the full potential of natural history collections and museums to serve society. After reviewing collections and types of museums, including local and global efforts, we discuss the value of specimens and the importance of observations. We then focus on mapping and modelling of museum data (including place-based approaches and discovery), and explore the main projects, platforms and databases enabling this. We also explore ways in which improved infrastructure will allow higher quality science and increased opportunities for interdisciplinary research and communication, as well as new uses of collections. Finally, we aim to improve relevant protocols for the long-term storage of specimens and tissues, ensuring proper connection with tomorrow's technologies and hence further increasing the relevance of natural history museums.

**AUTHOR COVER PAGE**

Article submission to PeerJ

Manuscript category: Literature Review Articles

Collection: "*Endless forms: Advances in evolutionary analyses of biodiversity*"

Article title:

# The Global Museum: natural history collections and the future of evolutionary biology and public education

Authors: Freek T. Bakker[1,*], Alexandre Antonelli[2,3,4,5,11], Julia Clarke[6], Joseph A. Cook[8], Scott V. Edwards[2,16], Per G.P. Ericson[10], Søren Faurby[3,4], Nuno Ferrand[12], Magnus Gelang[9], Rosemary G. Gillespie[13], Martin Irestedt[10], Kennet Lundin[3,9], Ellen Larsson[3,4], Pável Matos-Maraví[3,4,17], Johannes Müller[14], Ted von Proschwitz[9], George K. Roderick[13], Alexander Schliep[15], Niklas Wahlberg[7], John Wiedenhoeft[15] and Mari Källersjö[5]

[1]Wageningen University, Biosystematics Group, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.

[2]Harvard University, Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, 26 Oxford St., Cambridge, MA 02138, USA.

[3]Gothenburg Global Biodiversity Centre, Box 461, SE-405 30 Göteborg, Sweden.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

[4]University of Gothenburg, Department of Biological and Environmental Sciences, Box 461, 405 30 Göteborg, Sweden.

[5]Gothenburg Botanical Garden, Carl Skottsbergs gata 22A, SE-41319 Göteborg, Sweden.

[6]Jackson School of Geosciences, The University of Texas at Austin, 2275 Speedway Stop C9000, Austin, Texas, 78712, USA.

[7]Department of Biology, Lund University, Sölvegatan 37, 223 62 Lund, Sweden

[8]Museum of Southwestern Biology, Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, NM 87131, USA.

[9]Gothenburg Natural History Museum, Box 7283, SE-402 35, Göteborg, Sweden.

[10]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, PO Box 50007, Stockholm 10405, Sweden

[11]Royal Botanic Gardens, Kew, TW9 3AE, Richmond, Surrey, UK.

[12]Museu de História Natural e da Ciência, Universidade do Porto, 4099-002 Porto, Portugal

[13]Essig Museum of Entomology, Dept Environmental Science, Policy and Management, UC Berkeley CA 94720-3114, USA

[14]Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Invalidenstraße 43, 10115 Berlin, Germany

[15]University of Gothenburg, Department of Computer Science and Engineering, Rännvägen 6B, Göteborg, Sweden.

[16]Gothenburg Centre for Advanced Studies in Science and Technology, Chalmers University of Technology and University of Gothenburg, SE-412 96, Göteborg, Sweden.

[17]Institute of Entomology, Biology Centre CAS, 370 05 České Budějovice, Czech Republic.

* Author for correspondence. E-mail: freek.bakker@wur.nl

## Abstract

Natural history museums are unique spaces for interdisciplinary research and for educational innovation. Through extensive exhibits and public programming and by hosting rich communities of amateurs, students, and researchers at all stages of their careers, they provide a place-based window to focus on integration of science and discovery, as well as a locus for community engagement. At the same time, like a synthesis radio telescope, when joined together through emerging digital resources, the global community of museums (the 'Global Museum') is more than the sum of its parts, allowing insights and answers to diverse biological, environmental, and societal questions at the global scale, across eons of time, and spanning vast diversity across the Tree of Life. We argue that, whereas natural history collections and museums began with a focus on describing the diversity and peculiarities of species on Earth, they are now increasingly leveraged in new ways that significantly expand their impact and relevance. These new directions include the possibility to ask new, often interdisciplinary questions in basic and applied science; inform biomimetic design; and even provide solutions to climate change, global health and food security challenges. As institutions, they are incubators for cutting-edge research in biology and simultaneously protect core infrastructure for present and future societal needs. In this perspective, we discuss challenges to the realization of the full potential of natural history collections and museums to serve society. After reviewing collections and types of museums, including local and global efforts, we discuss the value of specimens and the importance of observations. We then focus on mapping and modelling of museum data (including place-based approaches and discovery), and explore the main projects, platforms and databases enabling this. We also explore ways in which improved infrastructure will allow higher quality science and increased opportunities for interdisciplinary research and communication, as well as new uses of collections. Finally, we aim to improve relevant protocols for the long-term storage of specimens and tissues, ensuring proper connection with tomorrow's technologies and hence further increasing the relevance of natural history museums.

**Keywords:** collections, field education, Global Museum, innovation-incubator, natural history, place-based, specimens, transcriptomics, epigenomics.

## Introduction

Natural history museums, which harbour extensive records of biological diversity, have always been meeting places for scientists, amateurs, and the public. By visiting a natural history museum and learning about nature, the lay citizen often tacitly endorses the information presented and considers it worthy of consideration and reflection. As such, museums are still considered trusted resources, at a time when many other institutions are bitterly mistrusted (*Foley, 2015*). Whereas this trustworthiness is true for most types of museum collections, natural history collections play a special role, in part because at the same time that they serve in public education, they are also actively used and curated by professional scientists to answer pressing problems in biology and beyond.

Museum biological collections are more than meets the eye. Each specimen harbours many kinds of data, such as information on locality and collection parameters, associated pathogens, biopolymers such as DNA and proteins, and metabolic compounds. This wealth of metadata across many specimens turns collections into powerful research tools, enabling scientists to test for historic environmental hypotheses and carry out diverse studies ranging from public health & safety (as cornerstones in studies of environmental health and epidemiology; *Suarez & Tsutsui*, 2004), biomimetic design –where naturally-occurring architectures and systems inspire technological innovation (*Jayaram & Full, 2016*; *Nirodi et al., 2018*), historical genomics (focussing on ancient alleles or past genotypes; *Bi et al. 2013; Besnard et al. 2014*), to global change (tracking shifts in phenotype across specimens through time; *Jones & Daehler, 2018*), something that a database of mere species observations cannot do. But natural history collections face challenges. They are in need of constant sustenance, funding, and curation.

## Survey Methodology

This paper aims at reviewing and conceptualising how natural history museums have had and could have a continuing and increasing role in enabling discovery in evolutionary biology and beyond. We also discuss their central role in training the next generation of innovative thinkers. The paper is the result of a three-day workshop on ‘The role of museums in modern evolutionary biology’ organized by Chalmers University of Technology and the University of Gothenburg (Sweden), under the auspices of the Gothenburg Centre for Advanced Studies (GoCAS), and held during 7-9 June, 2017. Participants, most of whom were based at natural history museums (broadly defined to also include botanical gardens; see below) as researchers or directors, were chosen to represent diverse areas of museum-based science. They were often based in Europe, but also the United States, South Africa and South America (Peru). After several sessions of brainstorming about the current relevance of museums and roadblocks to expanding impact, we collectively drafted an outline and all authors participated in the writing and direction of the paper.

## Collections & types of museums

We define natural history museums in a broad sense, as institutions containing diverse physical specimens, and sometime also including seed banks, substantial living, frozen, or dried tissue collections, and genomic data, among others. These collections include material from rapidly disappearing extant species, as well as extinct species, many of which are from the most inaccessible parts of the Earth. Collections may have innate, historic biases in taxonomic coverage and sampling design, which might need to be considered for their further development.

Whereas historically natural history museums, in particular those in Europe, have been linked to the colonialist enterprise, increasingly specimens are held in local museums and maintained by local scientists and students, and international specimen acquisition and study are conducted in strong partnership with these institutions. International agreements, such as the Nagoya protocol, rightly mandate such participation under the terms of Access and Benefit Sharing. Additionally, citizen science increasingly contributes to collections, which today are housed all over the world, and serve as gems of diverse global centres of cutting-edge research (see Fig. 1). Natural history museums may be located at universities, sometimes without exhibits, or may include public exhibits, such as typically occurs in national,  state or regional entities. In many cases, regional collections, and their exhibits, reflect and strengthen visibility, appreciation, identity and awareness of local culture and fauna & flora, therefore playing an important and confirmative role for the visiting public. Collections that span long periods of time reflect the history of science as well as changes in norms and values in society: *what* was sampled, *how* and *why*. This emphasis is especially visible in open, regional, collections on display. Regionality therefore, can be considered a strength of collections and fulfills an important role in sustaining regional pride in biodiversity. On the other hand, for many scientists, usually on a global research mission, regional collections could appear to be insufficient as it needs ‘global collections’ (see below, *The Global Museum*), for instance, to capture the full extent of evolutionary and biogeographic diversity of particular clades. Still, given the increasingly connected network of regional museums, the local depth provided by such regionality would be impossible to recover – if only for practical reasons - in a single, global museum, were it composed solely of physical specimens.

The distributed nature of the world's museum collections increases long term data security. Collections of natural objects will always require a specific physical location and this makes them vulnerable to theft, fire and water damage. For instance, it was thanks to this distributed nature that we still have duplicates of the tens of thousands of specimens of plants stored at the Berlin Herbarium when it was almost completely destroyed during World War II. The Butantã Museum in São Paulo had a world-renowned alcohol collection of 85,000 snakes and half a million Arachnids. After it was drastically destroyed by a fire in 2010 (*Phillips, 2010*), it became clear that there were virtually no duplicates for these two collections, either as specimens or tissues, in other museums. Along with these concrete recent examples of natural heritage loss, the infrastructure of many museums remains underfunded, exacerbating their vulnerability. A grim example is the Brazil National Museum in Rio de Janeiro, where a fire destroyed an estimated 90% of the collections in several divisions in September 2018 (*Phillips, 2018*).

Creating redundancies in collections, especially for extant species and genetic resource collections, is key to ensuring the longevity of these samples and associated data. Initiatives such as the Global Genome Biodiversity Network (GGBN; *Droege et al., 2016*) aim to collect, catalogue, and "democratise" genomic resources across global collections, covering 50,626 species (as of 18 march 2019). Although this enterprise represents an important step in the direction of a distributed collection, it will benefit from more coordination and financial support for data security and achieving complementarity and redundancy among collections. These aims are included in the mission of the pan-European Distributed System of Scientific Collections (www.dissco.eu) initiative, which aims to extensively digitize and database diverse specimens from across the museums of Europe. In comparison, in the US, a National Science Foundation-funded major digitization initiative, named *Integrated Digitized Biocollections* (iDigBio), aims at making data and images for millions of biological specimens available in electronic format "for the research community, government agencies, students, educators, and the general public" (*Page et al., 2015*). iDigBio serves as "the coordinating center for the national digitization effort" fostering partnerships and innovations, and developing extensive content. iDigBio is funded by grants from the NSF Advancing Digitization of Biodiversity Collections (ADBC) program. In contrast, DiSSCo started by securing government buy-in, with content being part of operational costs. Because it has become an established and recognized entity, governments can fund the infrastructure because they need the services, thereby providing DiSSCo with guaranteed political context. At the global level, the *Global Biodiversity Information Facility* (GBIF) is "an international network and research infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth." Being the main global database, a large proportion of its 1 billion records comprises observations rather than specimens (see below).

## The value of specimens and the importance of observations

All collections ultimately contain, or are dependent on, specimens. A specimen may consist of a complete organism (collected by naturalists over the past few centuries) or parts of a single individual organism. Increasingly, meta-data associated with the physical specimen - the 'extended specimen' (Webster, 2017) - add value and increase data richness through videos, sound recordings, information on habitat, and photographs. For example, for birds, the extended specimen may be comprised of records of the song, or recordings of behaviour of those organisms, prepared in a way that preserves them for the future. Bioacoustic tools provide unique collections that can include some of the last known evidence of extant species. Likewise, several films exist (e.g. https://www.youtube.com/watch?v=nAzqGn-LHCw) portraying the behaviour of now-extinct animal species, such as the Tasmanian tiger, the golden toad, and the Hawaiian crow. Museums currently host increasingly diverse collections, which, in addition to DNA and tissue banks, may be generated by core genomic facilities or imaging labs (isotopic, x-ray computed tomography data [CT], scanning electron microscopy [SEM] images). Examples now include rich stores of high-resolution CT data generated from museum specimens, which allow investigators to look inside material in a largely non-destructive way. These require different storage resources from those that traditionally constitute museum infrastructures, namely large scale and secure long-term storage of image data. Integration of different data streams will allow bridging among disciplines and the involvement of fields underrepresented in natural history museums, such as engineering, biomedical sciences, and art. For instance, biomimetic design can benefit strongly from inspiration from natural history collections (examples from robotics see *Jayaram & Hull, 2016*), or solutions to global health or food security challenges can be based on exploration of natural history specimen collections (see Table 3 *Specimens and pathogens*).

Specimens are at the heart of the discovery process and technological advances are increasing the number and diversity of possible questions that can be addressed (e.g., *Schmitt et al. 2018*; see below). For instance, bone fragment identification using collagen barcoding was difficult to imagine before the rise of LC-MS technology, but *Welker et al. (2015)* used this to identify Palaeolithic fragments of mammal bones in France. Genomic analyses of single bone fragments can inform on the evolutionary and demographic history of our own species (e.g., *Slon et al., 2018*). Future technologies may include more advanced chemical, biochemical, isotope or micro-anatomical surveys, making maintenance of specimens even more critical because they connect all data streams and facilitate data interoperability. At the same time, maintaining specimens is key to repeatability – a core requirement of science.

Museum specimens and data are stored and made accessible for future generations, enabling them having broad reference and context. Continued field collecting secures future access to time series of specimens, collected over decades or even hundreds of years. These long-term archives provide valuable and unique information (Graham *et al*. 2004) on changes in the species composition in our environments and habitats, due to factors such as climate change, human-mediated nitrogen deposition, or other anthropogenic activities (*Meineke et al*. 2019a; 2019b). An example is a large survey and collection of marine invertebrates from the Swedish west coast from the 1920´s and 30´s conducted by the Gothenburg Natural History museum (GNM), in which the exact sample locations could be deduced using modern methods, and consequently revisited during a new survey in the 2000´s, revealing a 60% loss of biodiversity (*Obst et al. 2017*). Specimens collected by researchers 200 years in the past can be compared with contemporary (and future) sampling-- as long as these collections and institutions persist. In fact, the specimen can be seen as the outcome of a combination of genotype and past environmental change or conditions (e.g., *Holmes et al., 2017*), and a well-curated collection captures the variation in phenotype as well as genotype (see *Bi et al*. 2013; *Rowe et al*. 2011; *Staats et al.* 2013; *Ruane and Christopher, 2017*). For instance, *Cridland et al. (2018)* comparing SNP patterns from historic museum and living specimens of bees, could not only infer ‘rapid change’ in genetic composition of honey bees in California, but also identify historic genotypes in candidate genes possibly involved in adaptation to new niches.  As another, less-domesticated, example: reconstructing the shift to C4 photosynthesis in grasses could be conducted using DNA from a 100-year-old Malagasy herbarium specimen for which both its phylogenetic placement and the assessment of its ‘genetic make-up’ with regards C4 photosynthesis could be assessed (*Besnard et al., 2014*).Therefore, specimen collections can provide a powerful reference for functional genomics studies, in a world where phenotyping different genotypes, retrievable from the specimen, becomes more important.

Increasingly, growth of museum collections is the result of their increased relevance for ecological studies, in addition to input from taxonomically-focused collecting activities, linked to specific inventories and research projects. Whereas museum staff and associated researchers and students still undertake expeditions to increase collections and make them available for future generations, many collections now come from large scale ecological studies (e.g. NEON in the US). Specimen collections enable answers to a large number of other scientific questions, some of which have not yet been posed. The earliest museums facilitated interactions among scholar-travellers, to share observational data from across the planet and to help build the core of what would become natural history and modern evolutionary biology. Increasingly, museums are leveraging new data from their specimens, and this integration of data types allows training in techniques that bridge among disciplines, as well as the generation of data sets that are of relevance to disparate traditional fields such as engineering, biomedical sciences, and art. Today natural history museums serve increasingly as a *nexus* for work that disregards disciplinary boundaries and addresses questions we did not know to ask before (see Fig 2). Because collections provide the opportunity to rigorously examine diverse aspects of taxonomic, morphological, genetic, and chemical variation across vast temporal and spatial scales, they can help diverse scientists bridge the gaps between traditionally distinct disciplines. Museum spaces ideally are filled with students who learn to think anti-disciplinarily and appreciate the importance of the specimen. These spaces can therefore be considered ‘Innovation Incubators’ where a next generation of critical thinkers in biology and beyond will be trained.

A specimen constitutes a voucher, not only of the actual individual sampled at the time, but often also of its locality – including information about the soil and other biotic and abiotic conditions in which the specimen was collected (see below, the ‘holistic specimen’). *Troudet et al. (2018)* describe how over the past decades the proportion of specimen-based occurrences in GBIF has decreased from 68 to 18%, in favour of observation-based occurrences, mostly from contributions by citizen science efforts such as iNaturalist and eBird (see Table 1). This will have affected repeatability and ‘richness’ of systematics and evolutionary studies and the authors urge that “when impossible to secure, voucher specimens can be replaced by observation-based occurrences”, particularly when combined with ‘ancillary’ data such as recordings, pictures, DNA samples etc. In cases where ethical, conservational, or practical concerns exist, observation data instead of collected specimens provide additional (or occasionally substitute) contributions to our knowledge on where and when particular species occur. Recorded sightings, such as those from iNaturalist or e-Bird, include occurrences of diverse temporal range, and are pretty much the only observation-based data that are allowed in GBIF. In addition to such observations recorded in the field, however, collected specimens (when available) offer additional options for confirming or extending the original work using new analytical techniques. Similarly, sound recordings can be re-studied within the context of new evidence, leading to reciprocal illumination.

## Place-based discovery: different specimen data sets connecting to a location

As *Miller (2007)* puts it, “Places are not simply a semantic convenience. It is a meaningful lens for viewing the world because it is orderly with respect to geographic space.” As such, ‘place-based’ approaches in general focus on the characteristics and meaning of particular places as a fundamental starting point for a particular development or project. Especially in charity and community development work, place-based approaches aim at “giving power to the community in guiding systemic change” and therefore “being recognised as critical to responding effectively to certain community challenges” (www.qcoss.org.au/). In education, place-based approaches are thought to “identify, recover and increase the values of local cultural specificities” (*Monardo, 2018*).

Place-based learning and education is well developed (*Gruenewald and Smith, 2014*) and provides a context for local understanding and societal change. Natural history museums are well suited for hosting place-based activities, as well as making direct links between collections and associated data and societal activities and needs. The developing Island Digital Ecosystem Avatar (IDEA) project is one example (*Davies et al., 2016*), entailing “a systems ecology open science initiative to conduct the basic scientific research needed to build use-oriented simulations (avatars) of entire social-ecological systems.” Many specimens will have been collected and stored, for instance for DNA barcode reference libraries, for making this possible.

For biological collections and their associated and ensuing process of discovery, the place-based approach is relevant. ‘Place-based’ here does not ignore external, or global, evidence or connections to other geographical localities, and indeed seeks to understand how local information and processes are interconnected with those at a larger scale. Best practices for biological collections include a geographical reference for each item, as is included in the

Darwin Core (see https://dwc.tdwg.org/). When collected together, sets of items are necessarily place-based. In addition to standardised metadata directly associated with biological items (*Kissling et al., 2018*), many other types of information are place-based and can be collected at the same location and super-imposed on point specimen data. Examples include information about geology, ground and atmospheric chemistry, and archaeology. These, and other data layers – such as from GBIF (species occurrences), NCBI (DNA and amino acid sequences), Open Tree of Life (phylogenetic trees), Map of Life (abundance data), TraitBase (traits), GloBI (biotic interactions), see Table 1 - can be associated or combined with geographical location through a geographical information system (HOLOS), integrating across diverse data types and enabling testing hypotheses concerning causal impacts - the 'holistic specimen'. In a sense, this approach is comparable to correlative species distribution modelling (SDM) approaches such as using Maximum Entropy (*MaxEnt;*, *Philips & Dudik, 2008*), focussing on mostly abiotic and edaphic correlates.

In addition, the place-based approach can provide a baseline for understanding changes over time (*Billick et al., 2013*; *National Research Council, 2014*). Of particular interest here is the context of historical environmental change to which current changes can be evaluated and compared. An understanding of historical processes provides a means for predicting, or forecasting, how biological systems may respond to change in the near future. For instance, *Willis et al. (2008)* studied how climate change may affect phenology in some angiosperm species in Concord, Massachusetts. *Slingsby et al. (2017)* studied the interaction between fire and climate change on species diversity in the South African Cape Floristic Region, allowing modelling of future vegetation response.

In general, the additional value of place-based approach for scientific discovery includes the serendipity of collecting data over periods of change identified later, as well as the interaction of researchers sharing an interest in the same geographical location or region (*Michener et al., 2009*). Place-based initiatives associated with larger networks (see Table 1) can provide access and understanding to a diversity of communities, which is both democratic and allows broad participation in discovery. Examples of such initiatives include developing new natural history museums (Darwin Initiative, www.darwininitiative.org.uk) and DNA barcoding of local biodiversity (*van de Bank et al., 2008*; *Janzen & Hallwachs, 2016*). In addition to natural history museums, the benefits of a place-based approach are also shared with field stations, botanical gardens, and biological reserves (*National Research Council, 2014*).

## The Global Museum

As indicated above, many museums serve regional communities, and collections in such institutions usually reflect regional interests, fauna and flora, funding and research questions. Given that science is an international endeavour, the question can be asked as to whether evolutionary biology would be better served by enhanced ability to document and analyse patterns across regions, such as with the use of GBIF. For instance, for taxonomy, having the virtual, global, workbench of the Barcode Of Life Database BOLD (www.boldsystems.org; *Ratnasingham & Hebert, 2007*) allowed taxonomists globally to harmonise species delineations by collectively analyzing and interpreting DNA barcode patterns from global rather than regional data sets. Historically however, collections have been mainly curiosity-driven, emphasizing rarities (i.e. single individuals per species, a phenomenon which may actually be commonly-occurring), a pattern that still exists for certain organism groups and regions (*Novotný & Basset,* *2000*; *ter Steege et al., 2011*), and mostly the result of general collecting efforts. This *generalism* has resulted in an invaluable reference specimen base in today's museum collections, allowing comparison with living specimens, identifying relatives of medically- and economically-important species (for instance melon, *Sebastian et al., 2010*), or testing historic biological hypotheses (e.g. Délye et al. 2013). On the other hand, this patchy tradition of biological collecting has come at a cost to easily comparing organisms across large geographic regions or across temporal spans. For example, evolutionary biology would benefit from being able to analyze more common species represented in collections worldwide, because this would allow assessing phenotypic variation at much broader scales. In addition, assumptions about species ID based on morphology may be falsified by genetic data (*DeSalle et al.* 2005) - but also the reverse - revealing an unexpectedly high level of cryptic diversity in certain groups (e.g., *Hebert et al., 2004*). Such a mandate would require a concerted effort of museums globally to collect and archive specimens in a coordinated manner that would help document current biodiversity and variation of common species across the globe. Such an effort was originally planned to be conducted by NEON in the United States, but in some cases has fallen short of this goal (*Cook et al., 2016*). Other ventures include the above-mentioned BOLD (with iBOL extending its coverage) which holds 6.6M barcode records across 0.29M species, many of which are commonly-occurring. Future collections should continue to expand with specimens sampled widely across biodiversity, but in addition should amass commonly-occurring species, which can serve as environmental monitors, especially when sufficient metadata is also collected.

Museum communities are increasingly not confined by a single, local physical space but able to distribute their reach through innovations in technology. Databases and other online tools enable international access and an array of novel platforms facilitate participation of a broad swath of the public in discovery and documentation, from undergraduate classes to young children participating in aspects of the scientific process. Examples of such activities include encouraging children to make observations of butterflies in drawers, thereby building their sense of biodiversity. Another example is citizen science projects in which volunteers help in interpreting and digitizing information on old collection labels, as has been done for the Paris Herbarium (http://lesherbonautes.mnhn.fr/) and for brachiopod fossils at the Swedish Museum of Natural History.

From a Global Museum perspective, we may ask whether phenomena such as global change have been effectively documented in collections in the past so that we can use that ensemble of past collections to forecast future conditions. For instance, collections can help scientists document how C4 photosynthesizing plants have spread during recent decades as a response to the global increase in atmospheric CO2 concentration (*Besnard et al., 2014*), or how species extinctions may be overrepresented in particular clades or areas (e.g., *Ricklefs, 2006*). Such work would be impossible without having the integrated, properly digitised and databased platform that a Global Museum provides.

Large international data sharing initiatives (e.g. LifeWatch ERIC, GBIF, Encyclopedia of Life, BOLD and iBOL, see Table 1) allow access to collections by scientists and the public living far from privileged historic western centres for inquiry. For instance, GBIF alone provides access to now over a billion records of specimens and observations from around the world. iDigBio, GBIF, and the Atlas of Living Australia (ALA) and its affiliated atlasses are the world's largest and best-developed biodiversity data aggregators and mobilizers. As outlined above, DiSSCo is a developing initiative of major significance that will unify natural science collections in Europe. With increasingly distributed access to large datasets and online portals to large-scale computational resources, analysis of the “big data” of biodiversity records can also go global (see *Antonelli et al. (2015)* for an example in global angiosperm biogeography and speciation). Digitization of collections will be increasingly important in this respect; there are many valuable but undigitized collections residing in museums.

Citizen science can contribute significantly to building collections, as for instance seen in many entomological collections that grow these days by amateur entomologists donating their well curated collections for posterity. *eBird* and *iNaturalist* are excellent examples of connecting citizen with science in a highly-efficient manner (and then feeding into GBIF). The Gothenburg Natural History Museum (GNM) malacological collections have benefitted tremendously by citizen science efforts with devoted (‘advanced’) amateurs donating their often well-curated private collections. Based on these collections, which can be considered ‘environmental archives’, *Bolotov et al.* (2018) could infer from freshwater pearl mussel collections that morphology has changed in time correlated to environmental alteration and climate change. Based on historical and recent specimens from extensive geographical sampling, the authors concluded that the latter may well have accelerated the population decline in pearl mussels over the last 100 years. The study underlines the importance of preserving large collections (many individuals) to enable meaningful statistical analysis of morphological measurements. Table 4 lists another example from the GNM concerning garden slugs sent in by the general public.

## Further increasing the relevance of museums through digitization and imaging

To facilitate the coordination of collection and databasing efforts between museums – a necessity to achieve the Global Museum – it is vital to increase awareness of what knowledge is available, not only at regional museums but across museums globally. This is most easily achievable through digitization of the collections. Moreover, such digitization also opens the collections for a number of additional researchers interested in overall temporal or spatial patterns in biodiversity. iDigBio provides a good example of how digitization can be successful and provides outreach to a global user base. In its first 10 years, iDigBio has prioritized digitization of specimens that can drive collaborative research and answer specific biological questions. This specialization necessarily results in only a small fraction of available specimens being digitized.

A major question for the future is how the community should greatly expand the scope of digitized specimens. Just as haphazardly collected historical specimens often prove useful for research questions not envisioned during the collecting event, it is also likely that specimens digitized without a specific research question in mind will prove useful for answering scientific or societal questions, especially if digitized on a large scale. But digitizing ‘blindly’ must of course be balanced with the pressure of meagre resources; enabling citizen scientists to assist offers a good solution (*Rouhan et al., 2016;* see Table 4). Large amounts of metadata remain to be digitized and would generate knowledge on biogeography (geographic data of specimens), disease spread (genetic material from parasites), biological interactions (pollination data), phenology, or shifts in species distributions (*Suarez & Tsutsui, 2004*; *James & al., 2018*). New advances in image recognition through deep learning using neural networks are also likely to enable easy identification of many species, such as already implemented in the iNaturalist platform, and hence help digitization.

*Tracking specimen taxonomy.* For research on temporal patterns of global change, such as global warming studies or analyses of movement of hybrid zones, museum records provide a unique source of historical records. Because they are backed up by physical specimens, records can be identified to individual species irrespective of how many taxonomic changes have occurred since specimen collection, an endeavour that is often impossible if the records were purely observational. However, an important caveat of this effort is that it is rarely obvious from online databases on what taxonomical opinion each identification was based. Indeed, to capitalize on the inherent strength of specimen-based records a citation to the taxonomic revision behind each identification would be ideal, but is not always available.

*Combining specimen databases.* Another challenge related to digitization and to the value of the physical specimens is that specimens may be one of the only clear, if frequently underutilized, ways to identify duplicates between different databases. For example, this challenge is often encountered among mammalian fossils where several databases, including New and Old World Mammals (NOW) and PaleoDB, have been started independently. Together these databases completely document diversity for some groups, and out of ~1586 accepted species of carnivores and relatives (Carnivora, Hyanodonta and Oxyaenidae) in the fossil record, 1460 species are included in at least one of them (*Faurby, Werdelin, Antonelli,* unpublished). However, individually they are each highly incomplete (1121 species (6385 records) in NOW, 1040 species (6756) in PaleoDB). Analyses attempting to summarize the complete fossil record therefore needs to combine databases. For many uses, such as when estimating speciation or extinction rates (*Silvestro & al., 2014*), it would be highly desirable to resolve redundancy between databases, a challenging effort that could be made trivial if both databases contained the museum specimen ids for the records.

*Big clades, large collections.* Digitization will be extremely challenging for the most diverse taxonomic groups, such as typically large collections of insects, although there have been attempts to automate digitization of such groups (*Hudson & al., 2015*). Knowing the limits in geographical and temporal distribution and the size of the collection, especially for the most common groups, may suffice for further scientific analysis. Finer scale distributional data in such cases can easily be supplemented by citizen science initiatives (e.g iNaturalist), especially if associated with pictures or movies with smartphones that contain georeferenced and timestamped records that facilitate re-validation through inspection of the images by experts (see for instance iSpot; *Silvertown et al., 2015*).

Museums still harbour large amounts of undiscovered and undocumented information. The total number of specimens deposited in museum collections around the world may be as large as 1-2 billion (*Ariño, 2010*), and for herbaria an estimated 350 M specimens are known to be deposited in 3400 collections world-wide (*Soltis, 2007*). Moreover, statistical approaches to estimate the size of collections agreed in 2010 that less than 5% of the universe of natural history collections data is available in databases such as GBIF (*Ariño, 2010*), although this fraction has been decreasing, with the fast increase of observation data in GBIF (*Troudet et al.*, 2018). *Wilson (2003)* noted that the smaller the organism the more poorly known the group to which it belongs, exemplified by fungi, nematodes and microbes. For instance, a random selection of specimens collected in a tropical rain forest and deposited in jars at a natural history collection resulted in the description of almost 200 new species of ichneumonid parasitoid wasps to science (*Veijalainen et al., 2012*). *Bebber et al. (2010)* described a comparable case for angiosperm species, with an estimated 35,000 undescribed species already residing in herbarium collections. An abundance of undescribed species is only the tip of the iceberg on the amount of data undiscovered and undocumented in the world's museums.

In terms of species diversity, DNA barcoding reference libraries such as BOLD (*Ratnasingham & Hebert, 2007*) and UNITE (*Koljag & al., 2013*) provide a good framework against which the extent of diversity deposited at natural history collections can be measured. However, molecular approaches to identify hidden diversity remain debatable (e.g., *Brower, 2006*) but can be overcome in large clades such as Lepidoptera (*Hebert et al., 2010*). In UNITE fungal species hypotheses are generated and named, but also tagged with a citable digital object identifiers (DOIs) so it can be unambiguously communicated, allowing harmonisation of species concepts throughout communities. In comparison, BOLD allows barcode index numbers (BINs) to refer to barcode clusters that have not been yet described taxonomically.

Programs like the UK Darwin Initiative to train observers and scientists in countries rich in diversity but low in funding for conservation and science surveys can further support democratization of not just specimens and data but also the knowledge for performing analyses and conducting research. However, much more needs to be done in this area, especially in capacity building, infrastructural development, and task distribution.

*Reference collections.* A major need for collections worldwide is to develop basic molecular data associated with a given taxon. The key importance is the burgeoning use of metabarcoding in ecological studies can be anchored to museum specimens, and thus linked to the associated metadata. Many museums have embarked on such endeavors, for example in CSIRO, and efforts of multiple museums ideally come together in clade-based DNA barcode projects in BOLD.

To enhance the broader relevance of natural history museums it is also important to message effectively to industry and policy makers. In particular, the museum community should explore ways to use specimens to find novel ways to bridge the traditional chasm between the sciences, arts and humanities. Shared themes include place-based research and experiential learning, both encouraged in instructional efforts in science, technology, engineering and mathematics (STEM fields) and the arts. Both the arts and sciences depend on inspiration, creativity, and critical assessment, and museum specimens serve well as sources for both inspiration and fascination. While scientific education and research offer rigorous methods for testing hypotheses and creating new knowledge, integration of experiential art and humanities work can into science fosters non-traditional ways of exploring and messaging about our world (Balengee 2010). Natural history museums should continue their efforts to train scientists and artists to develop novel solutions to emerging problems, especially as we face an increasingly uncertain environmental future.

Efforts by the Global Museum to assemble collections that will fulfil their key roles in the future require facilitating international agreement and participation. Such a massive effort cannot remain the province of a relatively few marginally resourced programs. Identifying the answers to the most pressing questions facing society and our environment require fertile spaces for cultivating innovation in the context of training in knowledge of biodiversity. This task is impossible without museum spaces and collection resources. We cannot afford to 'waste' the potential of natural specimens due to degradation, improper storage, or disposal, especially in the light of rapid biodiversity loss. They need space-efficient, climate-controlled and pest-free spaces. Innovations in these areas are likely needed to accommodate collections in the long term (hundreds to thousands of years) and to deal with preservation issues that may be exacerbated by global climate change (including increasing frequency of extreme weathers, hurricanes, flooding).

## Public perception of natural history museums

We argue that natural history museums should be regarded as ‘Innovation Incubators’, places where ‘anti-disciplinary’ science is thriving by building bridges between otherwise or so far improbable disciplines, and scientists from various ‘disciplines’ meet, an *Academic Nexus of Integration*. Because natural history collections can facilitate examining diverse aspects of taxonomic, morphological, genetic, and chemical variation across vast temporal and spatial scales, they can help diverse scientists bridge the gaps between traditional disciplines. In places where this situation is not yet in place the way to get there would be to enable ready access to both collections and research facilities, an effort that has been highly successful, for example, under the European SYNTHESYS Access scheme for the last decade. The K-12 education project “Exploring California Biodiversity” at the University of California, Berkeley (http://gk12calbio.berkeley.edu; *Mitchell & Gillespie, 2007*), which takes grade school students and teachers into the field, provides an excellent example of natural history museum collections broadening access and opportunities for education. Efforts such as AIM-UP! combined the expertise of educators, curators, collection managers, database managers, and others in undergraduate education (*Cook et al., 2016*; *Lacey et al., 2017*).

In terms of public perception of natural history museums, it is important to safeguard their role in society and justify long-term funding by continuing outreach and engaging the general public by proper messaging, for instance by initiating citizen science projects. In some sectors of the US it is still clear that the public misunderstands the mission of museums and does not appreciate the need for continued responsible collecting. One recent example is the unwarranted overreaction against scientific collecting of a bird specimen from the Solomon Islands, information about which was placed on the web by well-meaning media directors at the American Museum of Natural History (*Johnson, 2018*). This sad event, which resulted in death threats and cyberharassment of the scientist involved, shows that many people see only the destructive, indeed amoral, aspects of collecting of individual specimens, and do not connect this act with the many beneficial services of museums to science and society. Additionally, the public in this case did not appreciate the relative insignificance of scientific collecting as an agent of species loss as compared with habitat loss and introduced or feral predators, such as house cats. In this case, the public seemed to place undue emphasis on the loss of life incurred by collection of a single bird, suggesting much more relevance of an emotional response rather than a scientific appraisal of the true impact of collecting data on species biology. Clearly, museum curators and scientists need to join forces in working proactively with the public to increase their awareness of, and appreciation for, the practice of rigorous biological sampling.

## Collecting for the future: integrated analysis of museum specimens for evolutionary biology

Museums need room to grow in targeted ways that will allow us to address scientific issues critical to looming societal issues such as emerging pathogens and food security (*Morrison et al., 2017*; *Schindel & Cook, 2018*) (see Box 2). Specimen-based field work should aim to preserve extensive sets of natural history material at a particular time and place that would represent multiple individuals of each species, multiple species per collecting locale, and multiple diverse aspects of individual specimens. For example, collection of mammals and their associated ectoparasites and digestive tracts has led to detailed understanding of co-evolution of hosts and parasites (*Cook et al,. 2017*) and can fuel future studies of the role of the microbiome in such processes (*Roggenbuck et al., 2014, Greiman et al., 2018*). Such holistic collection events can better capture the complex interactions of biotic communities and, if repeated, over time could provide key insights into changing conditions. Discussions should be held that address how we can best leverage collecting activities across the Global Museum and that planning should lead to a global effort to more rigorously inventory biodiversity.

Genomics is a key source of information and rapidly changing area in which the scope and potential of future applications are particularly promising. Nonetheless, there are several factors known to limit the utility of specimens for genomic analyses. For example, using ‘methylated spirit’ (methanol containing alcohol) instead of pure alcohol for field preservation of animal tissue can severely hamper retrieval of usable DNA later on (*Post & al., 1998*) (see Box 1). Heat treatment of plants, as typically applied in most historic herbarium collections, was found to lower genomic copy numbers but not cause significant miscoding lesions (*Bakker, 2015*; *Staats et al., 2011*). Conventional x-rays (as opposed to x-ray computed tomography with digital imaging) of mummies and bone or using pesticides on insect collections all negatively affect or destroy DNA (*Gotherstrom et al., 1995*). Use of formalin to preserve specimens limits extraction of usable DNA from both animal or plant tissues as it causes cross-links among DNA molecules, preventing PCR (*Ruane and Christopher, 2017*; *McGuire et al., 2018*).

For historic samples, significant progress in securing biopolymers has been made and museum and ancient genomics has attracted considerable interest, from researchers and industry (*Hofreiter & al., 2015*; *Lindqvist & Rajora, 2019*). Still, although some DNA sequencing technologies work well with degraded DNA, such as in herbarium DNA using Illumina sequencing (*Staats & al., 2013*; *Bakker et al., 2016*; *Hart et al., 2016*), single-molecule, ‘3rd generation’, genome sequencing will never be applicable for most museum-preserved specimens given the fragmented nature of their DNA. In contrast, for cryo-specimens preserved in liquid nitrogen in tissue banks, 3rd generation sequencing may well be highly successful. But even here, some of us have noted poor DNA retrieval from tissues collected with standard field-protocols and ultimately preserved cryogenically (S. Edwards, pers. obs.). For birds, best practices for genome sequencing may not include freezing in the short term, which can fragment DNA, but rather unfrozen archiving of blood, which will preserve the longest DNA fragments. On the other hand, whether or not EDTA or 95% EtOH was used for DNA sample storage can be important too for successful long read sequencing (MI, pers. obs.).

'Re-sequencing', i.e. sequencing and mapping reads against a related reference genome sequence, has been successful in museum plants, fungi and insects (*Bi et al., 2014*; *Staats et al., 2013*). For organisms with relatively small genome sizes, such as birds the price for re-sequencing a genome from a study skin has become so low that curators of bird collections may consider to actually requiring complete genome sequencing for tissue from old museum samples. In this case, all parties ideally would benefit, the user for having access to the specimen, the museum for putting a halt to further specimen deterioration (as the genome sequence has been generated), and the next user for have both specimen and genome sequence available. That said, it is difficult to predict how DNA extraction techniques may evolve, and perhaps require considerably lower tissue amounts to produce higher DNA yields, meaning that high-throughput DNA extraction without an immediate use (DNA banking) is not an obvious choice for museums.

Often scientists endeavour to see inside museum specimens. Previous approaches such as dissection or histology are invasive techniques that necessarily result in the destruction of other data. Although recent imaging techniques (diceCT; *Gignac et al., 2016*) enable largely non-destructive work on these questions in non-model organisms preserved in alcohol, they do not completely ameliorate data loss due to selectivity in field materials preserved. Specimen field preparations may include freezing fresh tissue for DNA, preserving skeletons and skins but removing most internal organs and muscles. The latter obviously limits the kinds and diversity of research that can ultimately be performed on such specimens. For example, the vocal organ of birds was often not collected in birds despite the perceived importance of bird song and other vocalization.  Some now broadly used imaging techniques (e.g. diceCT) have not been studied for their effects on DNA/RNA amplification from formalin or alcohol-preserved specimens, and it is unknown if they further inhibit downstream molecular work involving these specimens.

For most large multicellular organisms, it is challenging to collect large numbers of tissues. However,  more portions of an organism can be feasibly preserved before discarding tissues when making new collections, particularly of common, easily accessed species. For example, at the Museum of Comparative Zoology at Harvard University, a typical avian specimen is now accompanied by 7-10 cryovials filled with DNA- and RNA-ready tissues from different organs, as well as at least one tube of unfrozen but refrigerated blood for genome sequencing. Such sampling will no doubt pose space challenges for long-term storage (which could be partially solved through the use of space-efficient biobanks), but is essential for a deep understanding of the effects of anthropogenic change on biodiversity (*Schmitt et al., 2018*). Integrating new imaging techniques into museum work flows will increase documentation prior to destruction (e.g. for genomic work). For instance, the Thermal Age Web Tool (http://thermal-age.eu/) was developed to help collections managers and users to quantify the risks associated with destructive analysis of specimens, based on calculated ‘thermal ages’ (*Smith et al., 2003*). The Synthesis of Systematic Resources programme (see http://www.synthesys.info/joint-research-activities/) provides further recommendations for non-destructive sampling of museum specimens and decision analysis as to how to best sample specimens for genomic research.

Lower price points for acquisition of genome data and some imagining techniques makes defining these best practices more urgent.

## Conclusions

For hundreds of years, natural history museums around the world have provided the general public and scientists with numerous opportunities to learn more about our natural world. Taken together, this 'Global Museum' must be seen as one of the most valuable assets of modern society and culture, providing the material to address challenges facing humanity today – such as baseline information against which to test hypotheses of local and global environmental change – and a critical regional cultural touchstone for the public. Natural history museums can function as inter-disciplinary meeting places, or *innovation incubators*, where questions are addressed that we did not consider asking before. The core of these institutions are the specimens. To maximise their use, it is therefore imperative to carefully consider how to best sample, preserve, handle, and store specimens in ways that not only meet today's demands but also new, unforeseen needs. Viewing natural history museums as critical infrastructure for scientific inquiry and public understanding may help raise their profile and awareness, facilitating continued support.

Despite their immense value, natural history museums are facing grand challenges. Taxonomic expertise is decreasing for many organism groups or is not represented in the curation of some collections. Funding often relies on public sources and may be adversely affected by political and socio-economic changes, comprising the long-term continuity of a museum's activities. New international regulations on the collection, export and use of specimens for non-commercial and commercial purposes are now increasing administrative burdens and may prevent further development of collections. Cross-institutional, international coordination of secure data standards has not yet been fully realised. We urge scientists, citizens, and policy-makers to give natural history museums the proper recognition they deserve as data archives, innovation hubs, and *Academic Nexūs of Integration*.

## Acknowledgements

This paper is the result of a Workshop on ‘The role of museums in modern evolutionary biology Workshop’ in June 2017, organized by Chalmers University of Technology and the University of Gothenburg (Sweden), under the auspices of the Gothenburg Centre for Advanced Studies (GoCAS). We are very thankful to Karin Hårding, Gunnar Nyman, and Mattias Marklund for their continuous support and assistance in the program as GoCAS organizers.

## Funding

We acknowledge funding from the following sources: the Swedish Research Council (B0569601), the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013, ERC Grant Agreement 331024), the Swedish Foundation for Strategic Research, a Wallenberg Academy Fellowship, the Faculty of Sciences at the University of Gothenburg, the Wenner-Gren Foundations, and the David Rockefeller Center for Latin American Studies at Harvard University to A.A.; the Marie Sklodowska-Curie research fellowship (European Commission, project MARIPOSAS-704035) to P.M.-M.

## Figure legends

Figure 1. The centrality of natural history collections to evolutionary biology and public understanding. Users, contributors and stakeholders of natural history collections are indicated; yellow arrows represent data flow, green arrows the flow of specimens.

Figure 2. Natural history museum specimens drive the cycle between inspiration and innovation, which spans across exhibitions and collections.

Figure 3. Word cloud illustrating relative abundance of topics and themes covered in this study.

## References

**Alivisatos AP, Blaser MJ, Brodie EL, Chun M, Dangl JL, Donohue TJ, Dorrestein PC, Gilbert JA, Green JL, Jansson JK, Knight R, Maxon ME, McFall-Ngai MJ, Miller JF, Pollard KS, Ruby EG, Taha SA, Unified Microbiome Initiative Consortium. 2015.** A unified initiative to harness Earth's microbiomes. *Science* **350**:507–508 DOI: 10.1126/science.aac8480.

**Ariño AH. 2010.** Approaches to estimating the universe of natural history collections data. *Biodiversity Informatics* **7:**81–92.

**Bakker FT. 2015.** DNA sequences from plant herbarium tissue. In: Hörandl E, Appelhans M, eds. *Next-generation sequencing in plant systematics,* p271-284. International Association for Plant Taxonomy (IAPT)

**Bakker FT, Lei D, Yu J, Mohammadin S, Wei Z, Van de Kerke S, Gravendeel B, Nieuwenhuis M, Staats M, Alquezar-Planas DE, Holmer R. 2016.** Herbarium genomics: Plastome sequence assembly from a range of herbarium specimens using an terative organelle genome assembly (IOGA) pipeline. *Biological Journal of the Linnean Society.* **117:**33–43 http://dx.doi.org/10.1111/bij.12642.

**Balengee, B. 2010.** Malamp, the occurrence of deformities in amphibians, Arts Catalyst, London and Yorkshire Sculpture Park, Wakefield, England.

**Bank M van de, Olivier M, Duthoit S, Savolainen V. 2008.** A DNA barcode for the flora of the Kruger National Park (South Africa). South African Journal of Botany **74:**370–371 10.1016/j.sajb.2008.01.073.

**Bebber DP, Carine MA, Wood JRI, Wortley AH, Harris DJ, Prance GT, Davidse G, Paige J, Pennington TD, Robson NKB, Scotland RW. 2010.** Herbaria are a major frontier for species discovery. *Proceedings of the National Academy of Sciences of the United States of America* **107:**22169-22171.

**Berenbaum MR and Zan. 1998.** Chemical phenotype matching between a plant and its insect herbivore. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 13743–13748.

**Besnard G, Christin P-A, Malé P-JG, L'huillier E, Lauzeral C, Coissac E, Vorontsova MS. 2014.** From museums to genomics: old herbarium specimens shed light on a C3 to C4 transition. *Journal of Experimental Botany* DOI 10.1093/jxb/eru395

**Bi K, Linderoth T, Vanderpool D, Good JM, Nielsen R, Moritz C. 2013.** Unlocking the vault: next-generation museum population genomics. *Molecular ecology* **22:**6018-6032.

**Bolotov IN, Makhrov AA, Yu M, Gofarov, Aksenova OV, Aspholm PE, Bespalaya YV, Kabakov MB, Kolosova YS, Kondakov AV, Ofenböck T, Ostrovsky AN, Yu Popov I, von Proschwitz T, Rudzīte M, Rudzītis M, Sokolova SE, Valovirta I, Vikhrev IV, Vinarski MV & AA Zotin. 2018.** Climate Warming as a Possible Trigger of Keystone Mussel Population Decline in Oligotrophic Rivers at the Continental Scale. *Scientific Reports* **8:**35 DOI:10.1038/s41598-017-18873-y

**Brower AVZ. 2006.** Problems with DNA barcodes for species delimitation: “Ten species” of Astraptes fulgerator reassessed (Lepidoptera: Hesperiidae). *Systematics and Biodiversity* **4:**127-132

**Billick, I., I. Babb, B. Kloeppel, J. C. Leong, J. Hodder, J. Sanders, and H. Swain. 2013.** Field Stations and Marine Laboratories of the Future: A Strategic Vision. National Association of Marine Laboratories and Organization of Biological Field Stations. Available at http://www.obfs.org/fsml-future

**Colegate SM, Welsh SL, Gardner DR, Betz JM, and KE Panter. 2014.** Profiling of Dehydropyrrolizidine Alkaloids and their N-Oxides in Herbarium-Preserved Specimens of Amsinckia Species Using HPLC-esi(+)MS. *J Agric Food Chem.* **62(30):**7382–7392.

**Cook JA, Edwards SV, Lacey EA, Guralnick RP, Soltis PS, Soltis DE, Welch CK, Bell KC, Galbreath KE, Himes C, Allen JM, Heath TA, Carnaval AC, Cooper KL, Liu M, Hanken J, Ickert-Bond S. 2014.** Natural History Collections as Emerging Resources for Innovative Education. *BioScience* **64:**725–734.

**Cook, J.A., Galbreath, K.E., Bell, K.C., Campbell, M.L., Carrière, S., Colella, J.P., Dawson, N.G., Dunnum, J.L., Eckerlin, R.P., Greiman, S.E., Fedorov, V., Hass, G.M.S., Haukisalmi, V., Henttonen, H., Hope, A.G., Jackson, D., Jung, T., Koehler, A.V., Kinsella, M., Krejsa, D., Kutz, S.J., Liphardt, S., MacDonald, S.O., Malaney, J.L., Makarikov, A., Martin, J., McLean, B.S., Mulders, R., Nyamsuren, B., Talbot, S.L., Tkach, V.V., Tsvetkova, A., Toman, H.M., Waltari, E.C., Whitman, J.L., Hoberg, E.P. 2017.** The Beringian Coevolution Project: Holistic Collections of Mammals and Associated Parasites Reveal Novel Perspectives on Evolutionary and Environmental Change in the North. *Arctic Science* 3:585-617.

**Cook JA, Greiman S, Agosta RP, Anderson BS, Arbogast RJ, Baker W, Boeger RD, Bradley DR, Brooks R, Cole JR, Demboski AP, Dobson JL, Dunnum RP, Eckerlin J, Esselstyn K, Galbreath J, Hawdon H, Hoekstra S, Kutz J, Light L, Olson BD, Patterson JL, Patton AJ, Phillips E, Rickart DS, Rogers M, Siddall V, Tkach and EP Hoberg. 2016.** Transformational principles for NEON sampling of mammalian parasites and pathogens: a response to Springer and colleagues. *BioScience* **66:**917-919.

**Cook JA, Lacey EA, Ickert-Bond SM, Hoberg EP, Galbreath KE, Bell KC, Greiman SE, McLean BS, Edwards SV. 2016.** From Museum Cases to the Classroom: Emerging Opportunities for Specimen-Based Education. Pp 787-799 in Pavlinov I, ed.*"Aspects of Biodiversity", Volume 54, Archives of Zoological Museum of Moscow State University,* Moscow, Russia.

**Darwin Initiative**, https://www.gov.uk/government/groups/the-darwin-initiative

**Davies N, Field D, Gavaghan D, Holbrook SJ, Planes S, Troyer M, Bonsall M, Claudet J, Roderick G, Schmitt RJ. 2016.** Simulating social-ecological systems: the Island Digital Ecosystem Avatars (IDEA) consortium. *GigaScience* **5,** 1. http://doi.org/10.1186/s13742-016-0118-5

**Délye C, Deulvot C, Chauvel B. 2013.** DNA Analysis of Herbarium Specimens of the Grass Weed *Alopecurus myosuroides* Reveals Herbicide Resistance Pre-Dated Herbicides. *PLoS ONE* **8(10)**, e75117

**Droege G, Barker K, Seberg O, Coddington J, Benson E, Berendsohn WG, Bunk B, Butler C, Cawsey EM, Deck J, Döring M, Flemons P, Gemeinholzer B, Güntsch A, Hollowell T, Kelbert P, Kostadinov I, Kottmann R, Lawlor RT, Lyal C, Mackenzie-Dodds J, Meyer C, Mulcahy D, Nussbeck SY, Ó Tuama É, Orrell T, Petersen G, Robertson T, Söhngen C, Whitacre J, Wieczorek J, Yilmaz P, Zetzsche H, Zhang Y, Zhou X. 2016.** The Global Genome Biodiversity Network (GGBN) Data Standard specification. Database. baw125. DOI: -10.1093/database/baw125

**Dunnum JL, Yanagihara R, Johnson KM, Armien B, Batsaikhan N, Morgan L, Cook JA. 2017.** Biospecimen repositories and integrated databases as critical infrastructure for pathogen discovery and pathobiology research. *PloS Neglected Tropical Diseases*11: e0005133

**Foley, 2015.** the_Macroscope.org

**Gitzendanner MA, PS Soltis, GK-S Wong, BR Ruhfel, and DE Soltis. 2018.** Plastid phylogenomic analysis of green plants: A billion years of evolutionary history. *American Journal of Botany* 0(0): 1–11. doi:10.1002/ajb2.1048

**Gokhman D, Lavi E, Prüfer K, Fraga MF, Riancho JA, Kelso J, Pääbo S, Meshorer E, Carmel L. 2014.** Reconstructing the DNA Methylation Maps of the Neandertal and the Denisovan. Science 344: pp. 523-527 doi: 10.1126/science.1250368

**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(9), pp.497-503.

**Grayson P, Sin SYW, Sackton TB, Edwards SV. 2017.** Comparative genomics as a foundation for evo-devo studies in birds, Pages 11-46 in G. Shen, ed. Methods in Molecular Biology: Avian and Reptilian Developmental Biology. New York, Humana Press.

**Gruenewald, D.A. and Smith, G.A. eds. 2014.** Place-based education in the global age: Local diversity. Taylor and Francis Press, New York.

**Greiman, S., J. A. Cook, V. V. Tkach, E. P Hoberg, D. Menning, A. G. Hope, S. A Sonsthagen, and S. L.** Talbot. 2018. Museum metabarcoding: a novel method revealing gut helminth communities of small mammals across space and time. *International Journal of Parasitology.* 48:1061-1070.

**Hart ML, Forrest LL, Nicholls JA & CA Kidner. 2016.** Retrieval of hundreds of nuclear loci from herbarium specimens. *Taxon* **65(5)**: 1081–1092.

**Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004.** Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *PNAS* **101:** 14812-14817.

**Hebert PDN, deWaard JR, Landry J-F. 2010.** DNA barcodes for 1/1000 of the animal kingdom. Biol. Lett. (2010) 6, 359–362 doi:10.1098/rsbl.2009.0848

**Hess CM, Wang Z, and SV Edwards. 2007.** Evolutionary genetics of Carpodacus mexicanus, a recently colonized host of a bacterial pathogen, *Mycoplasma gallisepticum*. *Genetica* **129**:217-225.

**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, doi: 10.1002/bies.201400160

**Holmes MW, Hammond TT, Wogan GOU, Walsh RE, LaBarbera K, Wommack EA, Martins FM, Crawford JC, Mack KL, Bloch LM, MW Nachman. 2018.** Natural history collections as windows on evolutionary processes. Molecular Ecology https://doi.org/10.1111/mec.13529

**Holos**, Ecoengine, UC Berkeley, https://holos.berkeley.edu/about/

**Hudson LN, Blagoderov V, Heaton A, Holtzhausen P, Livermore L, Price BW, van der Walt S, Smith VS. 2015.** Inselect: Automating the Digitization of Natural History Collections. https://doi.org/10.1371/journal.pone.0143402

**IBOL** (International Barcode of Life), 2014. What is iBOL? Available at: http://ibol.org/about-us/what-is-ibol/

**Ingala MR, Simmons NB, Wultsch C, Krampis K, Speer KA, Perkins SL. 2018.** Comparing Microbiome Sampling Methods in a Wild Mammal: Fecal and Intestinal Samples Record Different Signals of Host Ecology, Evolution. Front. Microbiol., 01 May 2018 | https://doi.org/10.3389/fmicb.2018.00803

**James SA, Soltis PS, Belbin L, Chapman AD, Nelson G, Paul DL, Collins M. 2018.** Herbarium data: Global biodiversity and societal botanical needs for novel research: Global. *Applications in Plant Sciences*. John Wiley and Sons Inc. https://doi.org/10.1002/aps3.1024

**Janzen DH & Hallwachs W. 2016.** DNA barcoding the Lepidoptera inventory of a large complex tropical conserved wildland, Area de Conservacion Guanacaste, northwestern Costa Rica. *Genome* **59**(9):641-60. doi: 10.1139/gen-2016-0005.

**Jayaram K & Full RJ. 2016.** Cockroaches traverse crevices, crawl rapidly in confined spaces, and inspire a soft, legged robot. PNAS 113, 23 February 2016, Pages E950-E957

**Johnson KW. 2018.** The Ornithologist the Internet Called a Murderer. The New York Times, June 15, 2018.

**Jones CA, Daeler CC. 2018.** Herbarium specimens can reveal impacts of climate change on plant phenology; a review of methods and applications. PeerJ 6:e4576 https://doi.org/10.7717/peerj.4576

**Kissling WD, Ahumada JA, Bowser A, Fernandez M, Fernández N, García EA, Guralnick RP, Isaac NJB, Kelling S, Los W, McRae L, Mihoub JB, Obst M, Santamaria M, Skidmore AK, Williams KJ, Agosti D, Amariles D, Arvanitidis C, Bastin L, De Leo F, Egloff W, Elith J, Hobern D, Martin D, Pereira HM, Pesole G, Peterseil J, Saarenmaa H, Schigel D, Schmeller DS, Segata N, Turak E, Uhlir PF, Wee B, AR Hardisty. 2017.** Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological Reviews* https://doi.org/10.1111/brv.12359

**Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, Bates SB, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Griffith TGW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín RMP, Matheny PB, Nguen NH,**

**Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüssler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria T, Weiss M, Larsson K-H, 2013.** Towards a unified paradigm for sequence-based identification of fungi. *Molecular Ecology* **22:** 5271–5277.

**Lacey EA, Hammond TT, Walsh RE, Bell KC, Edwards SV, Ellwood ER, Guralnick R, Ickert‐Bond SM, Mast AR, McCormack JE, Monfils AK, Soltis PS, Soltis DE, Cook JA. 2017.** Climate change, collections and the classroom: using big data to tackle big problems. *Evo Edu Outreach* 10:2; doi 10.1186/s12052-017-0065-3

**Llamas B, Holland ML, Chen K, Cropley JE, Cooper A, Suter CM. 2012.** High-Resolution Analysis of Cytosine Methylation in Ancient DNA. PLoS ONE 7, e30226 34.

**Liede-Schumann S, Khanum R, Mumtaz AS, Gherghel I & Pahlevani A. 2016.** Going west – A subtropical lineage (Vincetoxicum, Apocynaceae: Asclepiadoideae) expanding into Europe. *Molecular Phylogenetics and Evolution* **94:** 436–446.

**Lindqvist C, Rajora OP (eds.). 2019.** Paleogenomics, Genome-Scale Analysis of Ancient DNA. Springer.

**Maldonado C, Barnes CJ, Cornett C, Holmfred E, Hansen SH, Persson C, Antonelli A, Rønsted N. 2016.** Phylogeny Predicts the Quantity of Antimalarial Alkaloids Within the Iconic Yellow Cinchona Bark (Rubiaceae: *Cinchona calisaya*). *Frontiers in Plant Science* **8:**391.

**Matos-Maraví P, Duarte Ritter C, Barnes CJ, Nielsen M, Olsson U, Wahlberg N, Marquina D, Sääksjärvi I, Antonelli A. 2019.** Biodiversity seen through the perspective of insects: 10 simple rules on methodological choices and experimental design for genomic studies. PeerJ Preprints 7:e26661v3; doi 10.7287/peerj.preprints.26661v3

**McGuire JA, Cotoras DD, O'Connell B, Lawalata SZS, Wang-Claypool CY, Stubbs A, Huang X, Wogan GOU, Hykin SM, Reilly SB, 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. (DOI:10.7717/peerj.4470).

**Meineke EK, Davies TJ, Daru BH, Davis CC. 2019a.** Biological collections for understanding biodiversity in the Anthropocene. Philosophical Transactions of the Royal Society B, 2019; 374: 20170386. doi: 10.1098/rstb.2017.0386.

**Meineke EK & Davies TJ. 2019b.** Museum specimens provide novel insights into changing plant-herbivore interactions. Philosophical Transactions of the Royal Society B, 2019; 374: 20170393. doi: 10.1098/rstb.2017.0393.

**Michener WK, Bildstein KL, McKee A, Parmenter RR, Hargrove WW, McClearn D, Stromberg M. 2009.** Biological field stations: research legacies and sites for serendipity. *BioScience* **59**, 300-310. http://dx.doi.org/10.1525/bio.2009.59.4.8

**Mitchell BJ, Gillespie RG. 2007.** Graduate Students Take to the Field in K–12 Education. *PLoSBiology* **5**(6): e162. doi:10.1371/journal.pbio.0050162

**Monardo B. 2019.** Innovation districts as turbines of smart strategy policies in US and EU. Boston and barcelona experience. In Smart Innovation, Systems and Technologies, 100, pp. 322-335 doi: 10.1007/978-3-319-92099-3_38

**Morrison SA, Sillett TS, Funk WC, Ghalambor CK, Rick TC. 2017.** Equipping the 22nd-Century Historical Ecologist. *Trends in Ecology & Evolution* **32**, 578-588.

**National Research Council. 2014.** Enhancing the Value and Sustainability of Field Stations and Marine Laboratories in the 21st Century The National Academies Press, Washington, DC. http://doi.org/10.17226/18806

**Nirody JA, Jinn J, Libby T, Lee TJ, Jusufi A, Hu DL, Full RJ. 2018.** Geckos Race Across the Water's Surface UsingMultiple Mechanisms. Current Biology28, 4046–4051 doi.org/10.1016/j.cub.2018.10.064

**Obst M, Vicario S, Lundin K, Berggren M, Karlsson A, Haines R, Williams A, Goble C, Mathew C, Güntsch A. 2017.** Marine long-term biodiversity assessment suggests loss of rare species in the Skagerrak and Kattegat region. *Marine Biodiversity*, doi: 10.1007/s12526-071-0749-5

**Page LM, MacFadden BJ, Fortes JA, Soltis PS, Riccardi G. 2015.** Digitization of biodiversity collections reveal biggest data on biodiversity. Bioscience 65(9): 841–842.

**Pauli JN, Newsome SD, Cook JA, Harrod C, Steffan SA, Baker C, Ben-David M, Bloom D, Brown G, Cerling T, Cicero C, Cook C, Dohm M, Eherlinger J, Federhen S, Frey B, Ghosh P, Graves G, Gropp R, Hobson K, Jordan C, O'Connell T, Pilaar Birch S, Poelen J, Ratnasingham S, Russell L, Stricker C, Uhen M, Yarnes C, Hayden B. 2017.** Opinion: Why we need a centralized repository for isotopic data. *Proceedings National Academy Sciences, USA* **114**: 2997-3001.

**Phillips SJ, Dudik M. 2008.** Modeling of species distributions with Max-Ent: new extensions and a comprehensive evaluation. *Ecography* **31:**161–175

**Phillips T. 2010.** Sao Paulo fire destroys one of the largest collections of dead snakes. The Guardian 16 May 2010.

**Phillips T. 2018.** Brazil National Museum: as much as 90% of collection destroyed in fire. The Guardian 4 September 2018.

**Portugal SJ, Cooper HJ, Zampronio CG, Wallace LL, Cassey P. 2010.** Can museum egg specimens be used for proteomic analyses? Proteome Science 8:40

**Post RJ, Flook PK, Millest AL. 1993.** Methods for the preservation of insects for DNA studies. *Biochemical Systematics and Ecology* **21:**85–92, https://doi.org/10.1016/0305-1978(93)90012-g

**von Proschwitz T. 1997.** *Arion lusitanicus* MABILLE and *A. rufus* (L.) in Sweden: A comparison of occurrence, spread and naturalization of two alien slug species. *Heldia* 4 Sonderheft 5: 137-138.

**Raffauf RF, von Reis Altschul S. 1968.** The Detection of Alkaloids in Herbarium Material. *Economic Botany* **22:**267-269

**Ramos-Madrigal J, Smith BD, Moreno-Mayar JV, Gopalakrishnan S, Ross-Ibarra J, Gilbert MTP, Wales N. 2016.** Genome Sequence of a 5,310-Year-Old Maize Cob Provides Insights into the Early Stages of Maize Domestication. *Current Biology* **26:** 3195-3201, https://doi.org/10.1016/j.cub.2016.09.036

**Ratnasingham & Hebert. 2007.** BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes*, doi: 10.1111/j.1471-8286.2006.01678.

**Ricklefs RE. 2006.** Global variation in the diversification rate of passerine birds. Ecology 87(10):2468-78

**Roggenbuck-Wedemayer M, Bærholm Schnell I, Blom N, Bælum J, Bertelsen MF, Sicheritz-Pontén T, Johannes Sørensen S, Gilbert MTP, Graves GR, Hansen LH. 2014.** The microbiome of New World vultures, *Nature Communications* **5** 5498 (2014), doi:10.1038/ncomms6498.

**Rouhan, Germinal, et al. 2016.** "The herbonauts website: Recruiting the general public to acquire the data from herbarium labels." Botanists of the Twenty-First Century: Roles, Challenges and Opportunities. United Nations Educational, Scientific and Cultural Organisation: 143-148.

**Rowe KC, Singhal S, Macmanes MD, Ayroles JF, Morelli TL, Rubidge EM, Bi KE, Moritz CC. 2011.** Museum genomics: low-cost and high-accuracy genetic data from historical specimens. Molecular Ecology Resources, 11(6), pp.1082-1092.

**Ruane S, Christopher A, C. 2017.** Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Molecular Ecology Resources* **17**, 1003-1008. doi:10.1111/1755-0998.12655.

**Schindel DE, Cook JA. 2018.** The next generation of natural history collections. PLoS Biol 16(7): e2006125. https://doi.org/10.1371/journal.pbio.2006125

**Schmitt CJ, Cook J, Zamudio K, Edwards SV. 2018.** Museum specimens of terrestrial vertebrates are sensitive indicators of environmental change in the Anthropocene. Phil. Trans. R. Soc. Lond. B. in review.

**Sebastian P, Schaefer H, Telford IRH, Renner SS. 2010.** Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. PNAS 2010 **107:** 14269–14273

**Shultz AJ, Baker AJ, Hill GE, Nolan PM, Edwards SV. 2016.** SNPs across time and space: population genomic signatures of founder events and epizootics in the House Finch (*Haemorhous mexicanus*). *Ecology and Evolution* **6**, 7475–7489.

**Silvertown J, Harvey M, Greenwood R, Dodd M, Rosewell J, Rebelo T, Ansine J, McConway K. 2015.** Crowdsourcing the identification of organisms: A case-study of iSpot. ZooKeys 480: 125–146. doi: 10.3897/zookeys.480.8803

**Silvestro D, Salamin N, Schnitzler J. 2014.** PyRate: a new program to estimate speciation and extinction rates from incomplete fossil data. *Meth. Ecol. Evol.* **5:**1126-1131 doi: 10.1111/2041-210X.12263

**Slingsby JA, Merow C, Aiello-Lammens M, Allsopp N, Hall S, Kilroy Mollmann H, Turner R, Wilson AM and JA Silander Jr. 2017.** Intensifying postfire weather and biological invasion drive species loss in a Mediterranean-type biodiversity hotspot. PNAS vol. **114**(18): 4697–4702

**Slon V, Mafessoni F, Vernot B, de Filippo C, Grote S, Viola B, Hajdinjak M, Peyrégne S, Nagel S, Brown S, Douka K, Higham T, Kozlikin MB, Shunkov MV, Derevianko AP, Kelso J, Meyer M, Prüfer K, Pääbo S. 2018.** The genome of the offspring of a Neanderthal mother and a Denisovan father. Nature https://doi.org/10.1038/s41586-018-0455-x

**Smith CI, Chamberlain AT, Riley MS, Stringer C, Collins M. 2003.** The thermal history of human fossils and the likelihood of successful DNA amplification. *Journal of Human Evolution* **45**, 203-217

**Smith, O. et al. 2014.** Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. Sci. Rep. 4, 5559

**Smith, R.W. et al. 2015.** Detection of cytosine methylation in ancient DNA from five Native American populations using bisulfite sequencing. PLoS ONE 10, e0125344 35.

**Staats M, Cuence A, Richardson JE, Vrielink-van Ginkel R, Petersen G, Seberg O, Bakker FT. 2011.** DNA damage in plant herbarium tissue. *PLoS ONE* **6:** *e28448*

**Staats M, Erkens RHJ, van de Vossenberg B, Wieringa JJ, Kraaijeveld K, Stielow B, Geml J, Richardson JE, Bakker FT. 2013.** Genomic Treasure Troves: Complete Genome Sequencing of Herbarium and Insect Museum Specimens. PLoS ONE 8(7): e69189. doi:10.1371/journal.pone.0069189

**Suarez AV, Tsutsui ND. 2004.** The Value of Museum Collections for Research and Society. Bioscience 54:66-74

**Troudet J, Vignes-Lebbe R, Grandcolas Ph, Legendre F. 2018.** The Increasing Disconnection of Primary Biodiversity Data from Specimens: How Does It Happen and How to Handle It?, *Systematic Biology*, **67**, Issue 6, 1 November, Pages 1110–1119, https://doi.org/10.1093/sysbio/syy044

**Veijalainen A, Wahlberg N, Broad GR, Erwin TL, Longino JT, Sääksjärvi, IE. 2012.** Unprecedented ichneumonid parasitoid wasp diversity in tropical forests. https://doi.org/10.1098/rspb.2012.1664

**Webster MS. 2017.** *The Extended Specimen: Emerging Frontiers in Collections-Based Ornithological Research.* Boca Raton, FL, USA: CRC Press.

**Welker F, Soressi M, Rendu W, Hublin J-J, Collins M. 2015.** Using ZooMS to identify fragmentary bone from the Late Middle/Early Upper Palaeolithic sequence of Les Cottés, France. Journal of Archaeological Science 54, pp. 279-286

**Wen J, Egan AN, Dikow RB, Zimmer EA. 2015.** Utility of transcriptome sequencing for phylogenetic inference and character evolution. in: Hörandl, E. & Appelhans, M.S. (eds.), Next-Generation Sequencing in Plant Systematics. Chapter 2, http://dx.doi.org/10.14630/000003© International Association for Plant Taxonomy (IAPT)

**Wilson EO. 2003.** The encyclopedia of life. *Trends in Ecology and Evolution* **18:** 77-80

**Yanagihara R, Gu SH, Song J-W. 2015.** Expanded host diversity and global distribution of hantaviruses: Implications for identifying and investigating previously unrecognized hantaviral diseases. In: Shapshak P, Sinnott JT, Somboonwit C, Kuhn J, eds. Global Virology. Identifying and Investigating Viral Diseases. New York: Springer-Verlag. 161±198

**Yates TL. et al. 2002.** The ecology and evolutionary history of an emergent disease: Hantavirus pulmonary syndrome. *Bioscience*, **52:** 989-998.

**Yoshida K, Burbano HA, Krause J, Thines M, Weigel D et al. 2014.** Mining Herbaria for Plant Pathogen Genomes: Back to the Future. *PLoS Pathog* **10**(4): e1004028, doi:10.1371/journal.ppat.1004028

**Yoshida K, Sasaki E, Kamoun S. 2015.** Computational analyses of ancient pathogen DNA from herbarium samples: Challenges and prospects. *Frontiers in Plant Science* **6**, Article number 771, 6p

**Zeyl E, Marcotegui P, Harris PD, Bakke TA, Bachmann L. 2012.** Nye parasitter fra gamle fisk / New parasites from old fish. Naturen 136, 274-280

**Zhang Q, Hill GE, Edwards SV, Backström N. 2014.** A house finch (*Haemorhous mexicanus*) spleen transcriptome reveals intra- and interspecific patterns of gene expression, alternative splicing and genetic diversity in passerines. *BMC Genomics* **15**:305.

## Table 1(on next page)

Main global and local aggregators of biodiversity data.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

**Table 1.** Main global and local aggregators of biodiversity data.

| Acronym | Mission | Funding; scope | Type of data | Volume of records (M) |
|---|---|---|---|---|
| **ADBC** | Advancing Digitization of Biodiversity Collections | US | | |
| **ALA** | Atlas of Living Australia. https://www.ala.org.au/ | Australia | Observations, specimens | 84.8 |
| **BOLD** | Barcode of Life Database | Canada; global | | 6.8 |
| **DiSSCo** | Distributed System of Scientific Collections; digitization and databasing of european specimen collections | Europe | Specimens | 1500 |
| **eBird** | Citizen science: the world largest biodiversity-related citizen science project, gathering information on bird sightings, archive it, and “freely share it to power new data-driven approaches to science, conservation and education.” https://ebird.org/home | Global | Observations | 100 ‘yearly’ |
| **EOL** | Encyclopedia of Life; Global access to knowledge about life on Earth | Australia, Egypt, US; global | Species descriptions | <1,9 |
| **GBIF** | Global Biodiversity Information Facility | Global | Observations, specimens | 1300 |

| | | | | |
|---|---|---|---|---|
| **GGBN** | Global Genome Biodiversity Network | Global | DNA, tissues, environmental samples | 3.8 |
| | | | | |
| **GloBI** | Global Biotic Interactions; species interaction data by combining existing open datasets. https://www.globalbioticinteractions.org/ | US; global | Species interaction data e.g., predator-prey, pollinator-plant, pathogen-host, parasite-host | >0.7 |
| **HOLOS** | Berkeley Ecoinformatics Engine: accessing and visualizing integrated biological and environmental datasets to address questions of global change biology. https://holos.berkeley.edu/ | US; global | Different kinds of biological and environmental datasets | n.a. |
| **iBOL** | International Barcode of Life; extending BOLD’s coverage. iBOL’s forthcoming BIOSCAN will activate a biomonitoring system for half the world’s ecoregions, metabarcoding assemblages and studying species interactions from 2,500 sites. iBOL.org | Canada; global | DNA barcodes and metadata | see BOLD |
| **IDEA** | Island Digital Ecosystem Avatar; place-based systems ecology for building simulations of social-ecological systems | US; Moorea | Specimens, observations | ? |

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

| | | | | |
|---|---|---|---|---|
| **iDigBio** | Integrated Digitized Biocollections; digitisation and databasing of US specimen collections | US | Specimens | 117.5 |
| **iNaturalist** | Citizen science: one of the world's most popular nature apps, sharing observations globally; https://www.inaturalist.org/ | US; global | Observations | <1 |
| **LifeWatch** | Biodiversity research, -management and -conservation priority setting | Europe | Research tools | n.a. |
| **iSpot** | Citizen science: experts helping citizen community to identify its wildlife observations. https://www.ispotnature.org/ | UK; global | Species identifications | 0.030 |
| **MoL** | Map of Life; providing species range and dynamics information and species lists for any geographic area. https://mol.org/ | Global | Occurrences, observations | 8.8 |
| **NCBI** | National Center for Biotechnology Information | Global | Nucleotide and amino acid sequences; genome annotations | 0.37 species covered |
| **NEON** | National Ecological Observatory Network; continental-scale environmental data, infrastructure for research, educational tools to work with large data. https://www.neonscience.org/ | US | Observations | ? |

| | | | | |
|---|---|---|---|---|
| **OToL** | Open Tree of Life; construct a comprehensive, dynamic and digitally-available tree of life by synthesizing published phylogenetic trees along with taxonomic data. https://tree.opentreeoflife.org/ | US; global | Phylogenetic trees and taxonomies | 2.6<br>OTU's in taxonomy |
| **Traitbase** | Ecological species characteristics, individual level species information. https://traitbase.info/whatis | Spain; global | Specific characteristics e.g. body size, diet or fecundity | ? |

## Table 2(on next page)

Specimens and best curation practices for the future.

**Table 2.** Specimens and best curation practices for the future.

*Collecting.* Recommendations for best preservation techniques for new specimens during field collection are as important as the final storage conditions for improving specimen long-term utility for genomics (*Matos-Maraví et al., 2019*). Documenting treatment practices is also key to facilitating future analyses enabled by as yet undiscovered technologies. The plethora of technological uses of museum specimens calls for a re-evaluation of how specimens are preserved. For centuries, plants have been pressed, animals mounted, marine specimens ethanol- or formalin-fixed and fungi dried. Although these standard preservation methods should still continue, if only because they constitute the bulk of biological collections thus far and have a proven track record of fostering discovery, whenever possible researchers should try to sample additional types of specimen parts, and organs and meta-data.

*Storing.* More studies need to be undertaken to improve relevant protocols for the long-term storage of specimens and tissues. Like digitization, banking of genetic resources by museums is an area of rapid innovation, particularly as next-generation sequencing methods have become more common. As museum tissue collections are accessed more frequently for genome projects, it has also become clear that the preservation standards and types of tissues preserved in museums are often inadequate for supporting the genomics enterprise. For example, a typical museum tissue sample from a bird, even if frozen in nitrogen hours after sacrifice in the field, yields DNA qualities and lengths unsuitable for 3rd generation long-read sequencing platforms such as PacBio and Oxford Nanopore. Such technologies rely on the use of long DNA fragments to start with, requiring specimen tissues be frozen immediately (within > 10-15 minutes) upon collection. Although it may be difficult to use liquid nitrogen in the field, one solution is to use so called Dry Shippers, which are dewars designed for safe transportation of tissues at the same temperature as liquid nitrogen but without actually containing any free liquid nitrogen. Such shippers are routinely allowed for transportation back to the lab by airlines and can often hold cold temperatures for ~3 weeks. Innovations in cryogenics are likely to transform collecting of genomic resources by museums in the future. Still, we can take comfort that even from dried, centuries-old traditional specimens, valuable genetic information can readily be obtained by hybrid-capture and other approaches (*Bi et al., 2013*; *Staats et al., 2013*).

At the same time, some genome sequencing strategies may alleviate the use of unwieldy cryogenic methods altogether, at least for accessing DNA. What seems to be most important is that DNA (and RNA) degradation is stopped as fast as possible after collection. For example, for birds, one way to achieve long DNA fragments for next-generation sequencing is to use unfrozen blood stored in Queen's lysis buffer, which has been used by ornithologists for decades and takes advantage of birds' nucleated red blood cells. Blood stored in this way, with minimal shaking that will cause shearing, is a reliable source of high molecular weight DNA and has been shown to yield better genome assemblies than DNA retrieved from museum-grade frozen tissues (S. Edwards, G. Bravo, pers. obs). An alternative could be to store collections in the field in DMSO, although this appears to prevent RNA sequencing (Irestedt, unp. data). Such best-practices for fieldwork and storage of genetic resources needs to be shared more widely and rapidly among the museum community. A useful platform for identifying both repositories and tissues for a wide range of taxa, often called biobanks, is provided by the Global Genomic Biodiversity Network mentioned above (GGBN see http://www.ggbn.org/ggbn_portal/). Below we outline various updates in storage of genetic resources for both animals and plants, highlighting issues facing museum curators and collection managers looking at the future:

*Transcriptomics and Epigenomics.* The ever-increasing number of genome sequences becoming available can be efficiently explored in terms of gene function through transcriptomics - the sequencing of all transcribed mRNA expressed at a certain time, physiological or developmental state for a particular tissue. In this way, the 1Kite (http://www.1kite.org/) and 1kp (https://sites.google.com/a/ualberta.ca/onekp/) projects, assembling 1000 transcriptomes of insects and plants, respectively, have expedited progress in both comparative and functional genomics and a better understanding of gene function across these clades (see for instance *Gitzendanner & al., 2018*). We would expect future specimens to play an increasing role in this respect, but only if we make sure to store our specimens in such a way that RNA is preserved, for instance by rapid cryogenic storage of use of RNA-friendly buffers like RNA-later. Additionally, a diversity of epigenomics approaches, such as methylation, Chip-seq and ATAC-seq, are emerging and potentially of great use to the field of evolutionary biology (*Grayson et al., 2017*). Epigenomics is already commonly applied in evolutionary studies of adaptation and development, and has recently made headway in examining epigenetic maps of extinct human and plant populations (*Llamas et al., 2012*; *Gokhman et al., 2014*; *Smith et al., 2014*; *2015*). Best practices for preservation of biomaterials for epigenetics has not yet been discussed in the literature, and will be an important additional consideration for museum curators in the future.

*Proteomics.* Given future technological developments, it is likely that proteomes will be determined and used for further functional studies across the Tree of Life. Additionally, collagen from bone tissues have been demonstrated to give species-level amino acid variation from specimens several millions of years old using a ZooMS approach (*Welker et al., 2015*). *Portugal et al. (2010)* report on proteomics in museum egg specimens and conclude that current approaches to proteomics in such specimens may be limited in coverage of the proteome. In any case, storing tissues in the best possible ways for proteomics, ideally, cryogenically in order to stop proteases from working, now ensures that such analyses can be conducted in the near future.

*Secondary metabolites.* Compounds such as alkaloids, glucosinolates, furanocoumarins, flavonoids or terpenes can be measured in museums tissues, especially from plants (*Berenbaum & Zan 1998*; *Colegate & al., 2014*; *Mithen et al., 2010*; *Raffauf & Von Reis Altschul, 1968*). Access to such metabolites enables testing historic biological hypotheses such as past responses to change in herbivores and climate; but also in case of invasive species and testing what secondary compounds may have driven invasive success in species such as *Vincetoxicum nigrum* (Asclepiadaceae) (*Liede-Schuman & al., 2016*).

*Stable isotopes.* Advancement of techniques for studying specimens include measuring of stable isotopes, allowing monitoring environmental/atmospheric changes over time, given a time series of museum specimens (reviewed in *Schmitt et al., 2018*). Because elements are not expected to degrade over time like biopolymers do, proper specimen storage conditions for isotope analysis is probably not critical. Limiting factors in such studies now is the availability of robust spatial sampling and time series of specimens for analysis. Properly tracking the vast quantities of data that are generated in these analyses directly to the specimens is also a challenge (*Pauli et al., 2017*).

*Non-standard tissues.* Classical botanical specimens comprise branches with leaves and fertile organs (flowers, fruits). For some vertebrates, such as birds and mammals, it is primarily the external morphology that is preserved in collections, whereas for amphibians, reptiles and fish, preservation in formalin and/or alcohol can yield DNA sequences with some effort (*Ruane et al., 2017*; *McGuire et al., 2018*). Many biobanks, particularly in US museums, now also include samples of frozen blood and tissue from vertebrates, typically heart, liver and muscle. However, many other parts of organisms not conventionally stored in museums are becoming important in the effort to monitor global change. For instance, there is great interest among diverse scientists in investigating the microbiome of species - including the bacteria present in the digestive system, and what roles they may have to the species' adaptations to the local environments (*Roggenbuck et al., 2014*; *Alvisatos et al., 2015*; *Ingala et al., 2018*). Similarly, tree bark may provide important information on chemical defenses of plants and hold implications for medical applications (*Maldonado et al., 2016*). Transcriptome studies in vertebrates are becoming increasingly common as a means to study species' ability to adapt to changing environments and anthropogenic change (e.g., *Zhang et al., 2014*) and are widely used in phylogenomics of invertebrates and plants (*Wen et al., 2015*). Such studies encourage careful sampling and preservation of whole organisms or all organs separately when appropriate.

## Table 3(on next page)

Specimens and Pathogens.

**Table 3.** Specimens and Pathogens.

Museum collections have provided fundamental infrastructure for identifying and mitigating emerging pathogens and zoonotic diseases by public health agencies (*Dunnum et al., 2017*) such as the Centers for Disease Control (CDC). When a new pathogen emerged in 1993 in the southwestern United States, rapidly killing 7 people, authorities needed to know: had this pathogen accidently been released into the wild, or was it a newly evolved pathogen, or had the virus always been present and simply not previously identified? Without the availability of specimen archives, scientists would not have been able to efficiently determine the pathogen source and answer these fundamental questions. Large tissue archives from the Museum of Southwestern Biology and other museums (*Yates et al., 2002*) enabled virologists to quickly identify that this previously unknown zoonotic hantavirus pathogen was hosted by the locally common deer mouse (*Peromyscus maniculatus*). This virus is apparently transmitted to humans through inhalation of viral infected feces. Subsequent emergence of other human cases elsewhere in the Americas (e.g., Chile, Argentina, Panama), but with higher mortality, mobilized other specimen-based research efforts that identified other new strains of hantaviruses in many rodent species over the next 2 decades and on multiple continents. More recently, museum specimens of other groups of mammals were screened, leading to a radically reshaped understanding of hantavirus evolution, ecology and host occurrence (*Yanagihara et al., 2015*). Not only were more rodent host species for these viruses identified, but numerous species of shrews, moles, and bat species worldwide also harbor their own hantaviruses. These specimens originated from multiple continents and the new discoveries significantly expand the potential risk to humanity of these pathogens and increase the burden on public health systems worldwide.

Other examples of pathogen outbreaks have examined historical progression of diseases often decades or centuries after the outbreak (*Schmitt et al., 2018*). The impact of an invasive bacterial pathogen from poultry on native songbirds has been studied using avian tissue samples collected just prior to the outbreak in the eastern US, albeit without any foreknowledge of the impending epizootic (*Hess et al., 2007*; *Shultz et al., 2016*). The Norwegian fish fauna is well documented in the scientific collection of the Natural History Museum, University of Oslo and collection material was screened for monogenean ectoparasitic flatworms of the genus *Gyrodactylus* that were (unintendedly) collected along with the fish (*Zeyl et al.* 2012). This yielded 13 flatworm species that are new to science, and an additional seven parasite species new to Norway. Three *Gyrodactylus* species were also recorded from new fish hosts, and in particular *G. pterygialis* appeared to have a broad range of host species, helping fishery biologists tremendously in understanding and managing fish populations. From plants, *Yoshida et al. (2014; 2015)* used potato herbarium in order to determine the genotype of the *Phytophtera infestans* strain that caused the great Irish potato famine in the 19th century (and concluded it was a 'one-off' type, never seen again). Herbarium DNA was also crucial in discovering ancient alleles in the grass *Alopecurus myosuroides* that are relevant to herbicide resistance but pre-dating human influence (*Délye et al., 2013*). Studies using genomic data of a 5,310-year-old maize cob (*Ramos-Madrigal et al., 2016*) have shown that our understanding of the process of domestication and early selection pressures needs adjusting.

# Table 4(on next page)

Citizen Science.

Citizen Science ‘Man-spread molluscs’ project: *send in the slugs from your garden to the museum for identification*.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

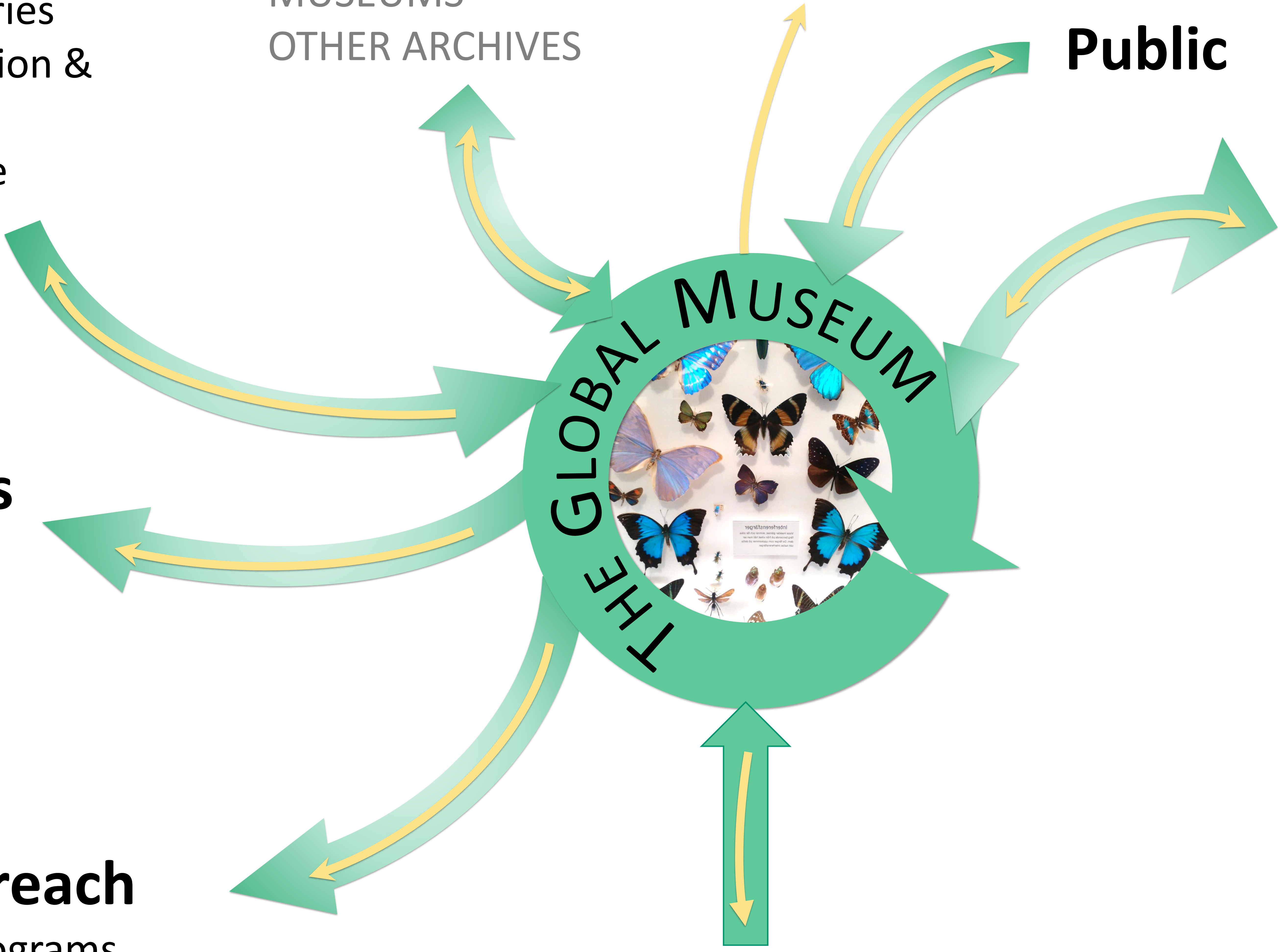
**Table 4.** Citizen Science ‘Man-spread molluscs’ project: *send in the slugs from your garden to the museum for identification*.

Since 1986 the Gothenburg Natural History Museum (GNM) has offered a slug-identification service to the public. The project was initiated as the invasive Spanish slug (*Arion vulgaris* Moquin-Tandon) began to spread rapidly over the country, prompting the need to establish a way to follow the spread and the colonisation process. As the slug spreads passively, mainly by the trade with ornamental plants and also with garden soil, it easily establishes in residential areas, where it mass-propagates and causes severe damage to vegetables and plants. A proper determination of these species requires dissection, a service offered by GNM. The project was advertised on TV, radio, in the newspapers and on museum web pages and GNM even offered to send out a transportation box which could easily be returned by mail. The response from the public was immense, and up to today GNM has received > 6000 samples with slugs from all over the country. Together with the box, GNM sent out a questionnaire asking information about first year of occurrence, possible way of introduction etc. After determination of the specimens the senders got information of species identity, and in case *A. vulgaris* was concerned also advice for control measurements. Hence the project has been beneficial for both the gardeners and the museum. The colonisation process could be followed in detail and much information on the biology and behaviour of the species, as well as on the garden fauna of snails and slugs, was obtained. The latter included several other invasive species, from different parts of the country, and the development of this fauna over more than four decades could be monitored both geographically and chronologically. Furthermore, as at least one specimen from each species in the samples was preserved inEtOH, extensive material is available for DNA analysis, which has proved highly useful as the taxonomy of the Spanish slugs is complicated, involving hybridisation with other native and introduced species (*von Proschwitz, 1997*).

## Figure 1(on next page)

The centrality of natural history collections to evolutionary biology and public understanding.

Users, contributors and stakeholders of natural history collections are indicated; yellow arrows represent data flow, green arrows the flow of specimens.



**Sample providers**

- Biodiversity Inventories
- Pathogen Identification & Monitoring
- Government Wildlife Agencies

UNIVERSITIES
MUSEUMS
OTHER ARCHIVES

WEBTOOLS
DATABASES

**Public**

**Science**

**Conservation Biology**

Population Status & Structure, Abundance, Bottlenecks, Invasives

**Ecology**

Distribution, Migration, Food Webs, Dispersal, Invasives, Stable Isotopes

**Emerging Pathogen Detection**

Identification, Monitoring
Host Switching, Range Expansion

**Evolution**

Response to Climate Change, Genome Evolution, Hybridization, Demography

**Policy initiatives**

- CITES
- CAFF
- NET-BIOME

**Educational outreach**

- Exhibits, Tours and Programs
- Curricula and Material for University Courses
- Web-accessible Biodiversity Data
- Internships (Field & Museum)

**Teaching & Research**

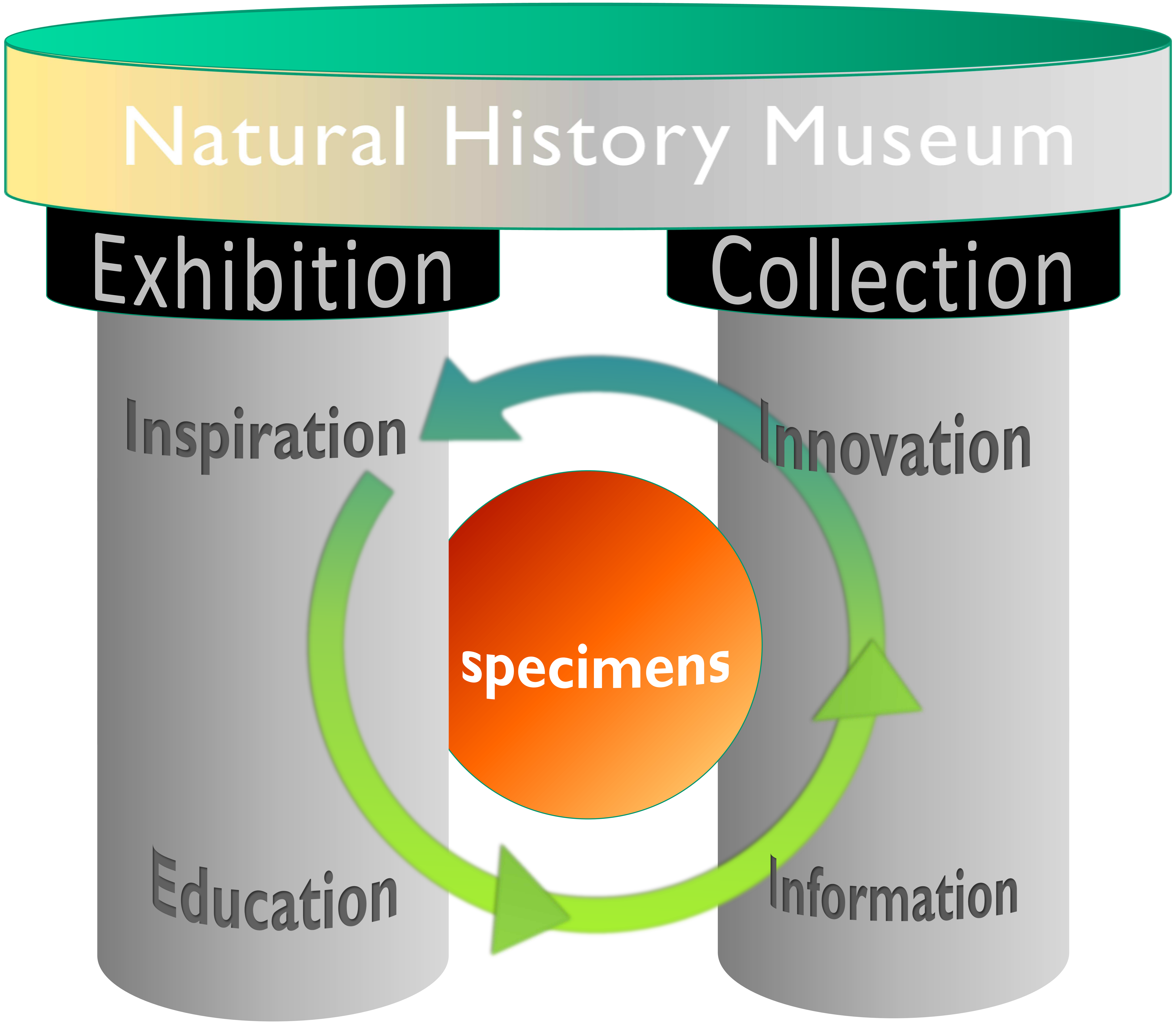
- Spatial and Temporal Perspectives
- Environmental Change
- Evolutionary Genomics
- Health and Economic Dimensions
- Systematics & Population Biology

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

## Figure 2(on next page)

Natural history museum specimens.

Natural history museum specimens drive the cycle between inspiration and innovation, which spans across exhibitions and collections.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27666v1 | CC BY 4.0 Open Access | rec: 19 Apr 2019, publ: 19 Apr 2019

## Figure 3(on next page)

Relative abundance of topics and themes.

Word cloud illustrating relative abundance of topics and themes covered in this study.


Museum
history
species
specimens
collection
specimen
collections
biodiversity
regional
place-based
citizen
help
discovery
long-term
preserved
climate
community
metadata
sciences
aspects
observations
students
GBIF
among
communities
organisms
taxonomic
impossible
specific
patterns
several
funding
recordings
individual
considered
past
recent
critical
allow
studies
best
living
used
tissue
techniques
sequencing
temporal
provide
knowledge
digitization
identify
Global
imaging
diverse
places
collecting
biology
facilitate
increasingly
digitized
education
disciplines
historical
understanding
development
Collections
efforts
addition
sampling
awareness
organism
institutions
changes
successful
building
based
need
include
tissues
public
global
researchers
DNA
genome
key
change
time
scientists
field
access
large
role
resources
future
scientific
environmental
groups
societal
hypotheses
museums
museum
local
new
work
serve
use
databases
world
loss
core
available
instance
effort
challenges
enable
years
sets
now
historic
especially
storage
spaces
Natural
genetic
questions
location
collected
biological
across
diversity
records
birds
ways
analysis
associated
activities
physical
society
increase
important
increasing
types
particular
without
evolutionary
provides
initiatives
international
generation
genomic
security
general
nature
science
data
natural
information
approach
approaches
geographical
therefore
occurrences
herbarium
example
research
History
barcode
systems
Examples