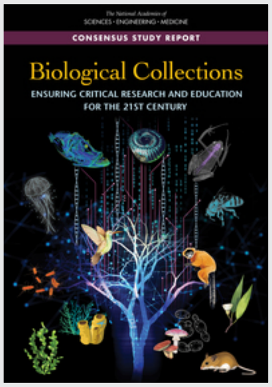


Biological Collections: Ensuring Critical Research and Education for the 21st Century (2020)

DETAILS

244 pages | 7 x 10 | PAPERBACK
ISBN 978-0-309-49853-1 | DOI 10.17226/25592

CONTRIBUTORS

Committee on Biological Collections: Their Past, Present, and Future Contributions and Options for Sustaining Them; Board on Life Sciences; Division on Earth and Life Studies; National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

National Academies of Sciences, Engineering, and Medicine 2020. *Biological Collections: Ensuring Critical Research and Education for the 21st Century*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25592.

Biological Collections

ENSURING CRITICAL RESEARCH AND EDUCATION FOR THE 21ST CENTURY

Committee on Biological Collections: Their Past, Present, and Future Contributions and Options for Sustaining Them

Board on Life Sciences

Division on Earth and Life Studies

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, NW Washington, DC 20001

This activity was supported by contracts between the National Academy of Sciences and the National Science Foundation's Division of Biological Infrastructure (Contract Number 10003964). Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project.

International Standard Book Number-13: 978-0-309-49853-1
International Standard Book Number-10: 0-309-49853-8
Digital Object Identifier: https://doi.org/10.17226/25592
Library of Congress Control Number: 2020947331

Additional copies of this publication are available from the National Academies Press, 500 Fifth Street, NW, Keck 360, Washington, DC 20001; (800) 624-6242 or (202) 334-3313; http://www.nap.edu.

Cover credits: Fresh kelp seaweed salad seafood vector illustration by Debahuti Bhattacharya on Shutterstock.com; set of cartoon underwater plants and sponges by Natali Snailcat at Shutterstock.com; *Escherichia coli* bacterium by Kateryna Kon on Shutterstock.com; watercolor vector hand-painted hummingbird by Elena Medvedeva on iStockphoto.com; DNA blue-colored double helix by KvitaJan on iStockphoto.com; modified handwork watercolor illustration of multicolored jellyfish by 4uda4ka on iStockphoto.com; computed tomography (CT) reconstruction of an angler, *Lophius piscatorius*, by Zachary Randall, Florida Museum of Natural History (reference UF 118531); 3D illustration of ammonite fossil by royaltystockphoto.com on Shutterstock.com; Micro-CT scan of a Burundi screeching frog, *Arthroleptis schubotzi*, by David C. Blackburn and Edward L. Stanley, Florida Museum of Natural History; modified hand-drawn honeybees by Val_Iva on iStockphoto.com; handwork watercolor illustration of red vari lemur by lenny777 on iStockphoto.com; watercolor vector hand-painted set with eucalyptus leaves by Elena Medvedeva on iStockphoto.com; illustration of yeast as part of the archives of Pearson Scott Foresman, donated to the Wikimedia Foundation; watercolor illustration of mushrooms by Darina_V on Shutterstock.com; illustration of a deer mouse (*Peromyscus maniculatus*) from See Pest & Lawn Solutions, Collegeville, Pennsylvania; digital tree on technology background representing the growth of modern-age digital media by monsitj on iStockphoto.com.

Printed in the United States of America

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2020. *Biological Collections: Ensuring Critical Research and Education for the 21st Century*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25592.

The National Academies of

SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, nongovernmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. John L. Anderson is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.nationalacademies.org**.

The National Academies of

SCIENCES • ENGINEERING • MEDICINE

Consensus Study Reports published by the National Academies of Sciences, Engineering, and Medicine document the evidence-based consensus on the study's statement of task by an authoring committee of experts. Reports typically include findings, conclusions, and recommendations based on information gathered by the committee and the committee's deliberations. Each report has been subjected to a rigorous and independent peer-review process and it represents the position of the National Academies on the statement of task.

Proceedings published by the National Academies of Sciences, Engineering, and Medicine chronicle the presentations and discussions at a workshop, symposium, or other event convened by the National Academies. The statements and opinions contained in proceedings are those of the participants and are not endorsed by other participants, the planning committee, or the National Academies.

For information about other products and activities of the National Academies, please visit www.nationalacademies.org/about/whatwedo.

COMMITTEE ON BIOLOGICAL COLLECTIONS: THEIR PAST, PRESENT, AND FUTURE CONTRIBUTIONS AND OPTIONS FOR SUSTAINING THEM

Co-Chairs

JAMES P. COLLINS, Arizona State University
SHIRLEY A. POMPONI, Florida Atlantic University

Members

ANDREW C. BENTLEY, The University of Kansas
RICK E. BORCHELT, Department of Energy
KYRIA BOUNDY-MILLS, University of California, Davis
JOSEPH A. COOK, The University of New Mexico
LYNN D. DIERKING, Oregon State University
SCOTT V. EDWARDS (NAS), Harvard University
MANZOUR H. HAZBÓN, American Type Culture Collection
TALIA S. KARIM, University of Colorado
GEORGE I. MATSUMOTO, Monterey Bay Aquarium Research Institute
PAMELA S. SOLTIS (NAS), University of Florida
BARBARA M. THIERS, New York Botanical Garden

Staff

AUDREY THÉVENON, Study Director
KEEGAN SAWYER, Senior Program Officer
JESSICA DE MOUY, Senior Program Assistant
ALYSSA R. FREDERICK, Christine Mirzayan Science & Technology Policy Fellow
MATTHEW ANDERSON, Financial Business Partner
FRAN SHARPLES, Director (*until February 2020*)
KAVITA BERGER, Director (*from July 2020*)

Consultant

ROBERT POOL, Editor

At this point in our history it is vitally important to acknowledge the fact that more and more of the species in biological collections will represent species, or certainly populations, that no longer exist as living organisms in nature. As scientists and as a society, we need to protect the specimens that we have, and to take special care with those we are collecting now. Equally important will be ongoing efforts to expand the types of living organisms we culture for research.

In many cases, museums and stock centers will, unfortunately, end up having the last remnants of species and populations that will never again exist on Earth. It's almost as if we had a few days to collect on another planet, and will never be there again.

In view of this situation, we need to think deeply and thoughtfully about the preservation of what we have, to collect and culture comprehensive specimens, ones for which material useful for genomic analysis is preserved, and then figure out how to keep our biological collections well maintained for as long as possible.

Peter H. Raven, President Emeritus, Missouri Botanical Garden

Preface

Biological collections are a critical component of the scientific infrastructure in the United States and globally. They advance scientific discovery and innovation, enrich education, connect communities to nature and science, and preserve Earth's biological heritage. Our nation's natural history and living stock collections enable research to improve health, food security, and national defense. Biological collections are used to reveal the history of life on Earth, study the impacts of humans on biodiversity, advance biomedical research, and develop improved crops, biocontrol agents, and pharmaceuticals.

Biological collections house living and preserved specimens that have a record of shedding light on the emergence and spread of pathogens and their hosts. Notably, the committee began working on this report before the coronavirus disease 2019 (COVID-19) pandemic started and finished it in the midst of the viral outbreak. Infectious diseases are a clear point at which living stock and natural history collections intersect in the service of society. COVID-19, for example, reminds us that pandemics and epidemics are not just ancient events, but under the right circumstances, new pathogens can emerge and cause great harm to modern societies. Biological collections provide the specimens needed to understand how infectious diseases emerge and how they might be mitigated before reaching the destructive level of the modern-day COVID-19 pandemic.

The ability to store, access, and use collections has significantly improved with new methods of automation, preservation, information extraction, data integration, and related technologies. Yet, despite the rich history of research, discovery, learning, and innovation made possible by biological collections, the infrastructure that supports them and makes them accessible deserves to be valued and appreciated much more than it is.

The biological collections community has produced many discerning and detailed reports on the needs, capabilities, and promise of biological collections. This Consensus Study Report echoes the findings of preceding publications while bringing new insights and a fresh perspective on ways to maintain, enhance, and expand the full portfolio of resources and assets that reside in biological collections. The report also reminds us that biological collections are part of the world's scientific infrastructure. Sustaining the priceless biological collections that are our heritage and our legacy is urgent if we are to continue to be able to address world-class scientific questions that depend on these kinds of collections, foster innovation, and support educational needs, now and in the future.

We extend our gratitude to the many experts who taught us about the range of challenges and accomplishments of biological collections. Their knowledge and insight through webinars, in-person presentations, and written comments sent through the project website stimulated rich discussion and enhanced the quality of the report. We also thank the external reviewers of the report for helping us to improve its accuracy. This report would not have been possible without the exceptional contributions of the National Academies of Sciences,

Engineering, and Medicine. Our committee is grateful to Audrey Thévenon, our study director, and Keegan Sawyer, senior program officer, for their guidance, dedication, and perseverance. Jessica De Mouy provided exemplary behind-the-scenes technical and logistical support for all of the committee's activities. Robert Pool substantially improved the language and format in our report.

The committee was fortunate to have a diverse and knowledgeable membership. The expertise, perspective, and dedication of the committee members cannot be overstated. We extend a special thank you to our colleagues on the committee who worked tirelessly to thoughtfully and carefully review a large amount of information and prepare this Consensus Study Report. It was an honor and privilege to work with all of them.

We hope that the committee's recommendations will provide inspiration and an evidence-based framework to build and support the nation's biological collections, which are crucial contributors to our capacity for discovery, innovation, and competitiveness now and for future generations.

James P. Collins and Shirley A. Pomponi, *Co-Chairs*
Committee on Biological Collections:
Their Past, Present, and Future Contributions and Options for Sustaining Them

Reviewers

This Consensus Study Report was reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise. The purpose of this independent review is to provide candid and critical comments that will assist the National Academies of Sciences, Engineering, and Medicine in making each published report as sound as possible and to ensure that it meets the institutional standards for quality, objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their review of this report:

MEGAN BANG, Spencer Foundation
ANN BARTUSKA, Resources for the Future
PHILIPPE DESMETH, Belgian Federal Science Policy Office
SARAH GEORGE, The University of Utah
GRETCHEN B. JORDAN, University of Maryland
PETER M. KAREIVA, University of California, Los Angeles
MICHELLE KOO, University of California, Berkeley
JAY LABOV, National Academies of Sciences, Engineering, and Medicine (retired)
MICHAEL LOMAS, Bigelow Laboratory for Ocean Sciences
KEVIN McCLUSKEY, Kansas State University
MURIEL POSTON, Pitzer College
FLOYD SHOCKLEY, Smithsonian Institution
SUSAN SINGER, Rollins College
JOSHUA TEWKSBURY, University of Colorado Boulder

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations of this report nor did they see the final draft before its release. The review of this report was overseen by **PETER H. RAVEN,** Missouri Botanical Garden, and **JOEL CRACRAFT,** American Museum of Natural History. They were responsible for making certain that an independent examination of this report was carried out in accordance with the standards of the National Academies and that all review comments were carefully considered. Responsibility for the final content rests entirely with the authoring committee and the National Academies.

Contents

Summary

For centuries, scientists have sought and collected different types of organisms to learn more about their forms, functions, origins, distributions, and evolution. Pooling and conserving these organisms into *biological collections*—systematized repositories of life in all of its many forms—is a cornerstone of quality research and education in many areas of science and innovation (see Box S-1).

Biological collections produce a wide range of benefits for science and education in the United States and the global community. Biological collections stand alone in providing the temporal, spatial, and taxonomic sampling of our natural heritage, preventing loss of knowledge about life on Earth. They support research on basic biological structures and processes and deepen our understanding of evolution, biodiversity, and global environmental change.

The health of biological collections—and, ultimately, of the scientific research that relies on them—is dependent on the underlying infrastructure that assembles, maintains, and provides access to these collections. Unfortunately, the sustainability of the nation's biological collections is under threat. The causes are many, including a general lack of understanding of their value and their contributions to research and education, a lack of appreciation for what is required to maintain them effectively, and inadequate coordination and interconnection among and between collections. It is easy to overlook the importance of infrastructure. When everything is functioning smoothly, infrastructure—whether it is the facilities of a university, the computers and transmission devices underlying the Internet, or the air traffic control system responsible for air travel—tends to be taken for granted. The same is true of the nation's biological collections. Without necessary changes in support and organization, the prior and current investments in time, money, and staff resources for building the nation's biological collections will be diminished, and their immense potential in supporting science, innovation, and education in the United States and elsewhere will be severely limited.

Recognizing the importance and the vulnerabilities of the nation's biological collections, the National Science Foundation (NSF) has endeavored to provide broad financial support through its Division of Biological Infrastructure within the Directorate for Biological Sciences. However, NSF welcomes guidance on a wide range of questions regarding long-term sustainability, including questions about operational structures, policies, and social cultures that could provide momentum to maintain and grow biological collections. For this reason, NSF asked the National Academies of Sciences, Engineering, and Medicine (the National Academies) to address the following:

- explore the contributions of biological collections of all sizes and institutional types to research and education;

BOX S-1
What Are Biological Collections?

Biological collections typically consist of organisms (specimens) and their associated biological material, such as preserved tissue and DNA, along with data—digital and analog (such as handwritten field notes)—that are linked to each specimen. Non-living specimens, which include organisms preserved by scientists and naturally preserved remains, such as fossils, are commonly referred to as natural history collections. Living specimens include research and model organisms that are grown and maintained in genetic stock centers, germplasm repositories, or living biodiversity collections. The defining trait of these different types of collections is that they capture aspects of the living world in such a way that it can be intensively studied and understood through time.

FIGURE S-1 Examples of biological collections in the United States. (A) spider in amber, University of Colorado Museum of Natural History Paleontology Section; (B) bats, Museum of Southwestern Biology, The University of New Mexico; (C) *Fusarium graminearum*, Fungal Genetics Stock Center, Kansas State University; (D) Xenopus, The National Xenopus Resource, Marine Biological Laboratory; (E) various herbarium specimens, New York Botanical Garden C.V. Starr Virtual Herbarium; (F) Charles Doe egg collection, Florida Museum of Natural History; (G) Ichthyology Cleared and Stained specimens in jars, University of Kansas Biodiversity Institute & Natural History Museum; (H) bacterial strain on petri dishes, American Type Culture Collection.

- envision future innovative ways in which biological collections can be used to further advance science;
- outline the critical challenges to and needs for their use and maintenance, including the quality control challenges faced by living stock collections; and
- suggest a range of long-term strategies that could be used for their sustained support.

The full Statement of Task for the study is provided in Appendix A. In responding to the Statement of Task, the committee considered two broad categories of biological collections: (1) non-living organisms, also referred to as natural history collections; and (2) living organisms, including research and model organisms.[1] In that regard, this report is the first of its kind. The committee acknowledges that living collections and natural history collections have distinct purposes and needs, but the committee also found that there are many opportunities for these communities to learn from one another and collaborate. Throughout the report, the committee highlights some of these potential synergies and intersections (e.g., digital genetic data, extended specimen information) as well as key distinctions (e.g., business strategies, quality control). This report is not an exhaustive compendium of every issue; rather, it focuses on challenges and paths forward for the biological collections community to work toward a common vision.

THE VALUE OF TODAY'S—AND TOMORROW'S—BIOLOGICAL COLLECTIONS

Biological collections are a critical part of the nation's science and innovation infrastructure. Although the number and extent of biological collections are unknown, scientists estimate that 800 million to 1 billion specimens are housed in U.S. natural history collections. Those, combined with living stock collections, which continually propagate and multiply organisms for research, result in a total number of U.S. biological specimens that undoubtedly exceeds 1 billion. The specimens are increasingly accompanied by a rich complement of additional biological material and data that are being used to generate new insights about life on Earth and to open new avenues of inquiry in almost every field of science, medicine, and technology.

Traditionally biological collections have been most heavily utilized by researchers trying to classify and understand the origins of biodiversity, including terrestrial and marine species as well as microbes. They provide the foundation

[1] NSF asked that these tasks be addressed for "living stocks (organisms) and preserved repositories of biodiversity specimens and materials" (i.e., natural history collections) that receive, or are eligible to receive, support for infrastructure or digitization from the NSF Division of Biological Infrastructure. This report does not explicitly address living collections in zoos, aquaria, or botanical gardens; biobanks or repositories of human tissues; or anthropological and geological collections (excluding fossils). This report also does not cover biological collections owned by federal agencies.

for scientific knowledge about how past and present organisms are interconnected, and the ways in which their physical and genetic characteristics change over time and space. However, specimens and their associated data—from genetic and molecular signatures to digital label data and images—also serve as source material for discovery and hypothesis-driven research. Numerous publications have documented how biological collections underpin basic discovery science. For example, the fruit fly *Drosophila melanogaster* has been used as a model organism for genetic research since Thomas Hunt Morgan used it to elucidate the role that chromosomes play in heredity, for which he was awarded the 1933 Nobel Prize. The discovery of the enzyme *Taq* polymerase in a bacteria strain deposited in a living stock collection led to the advancement and accessibility of next-generation sequencing technologies that rapidly transformed life science research by providing the ability to rapidly analyze and profile genomes. The development of the revolutionary genome-editing technique known as CRISPR (clustered regularly interspaced short palindromic repeats), which vastly expanded the genetic resources available in living collections and advanced the applications of biotechnology in medicine, agriculture, and conservation, was also, in part, the result of research on materials sourced from living microbe collections.

Biological collections also support much of the applied research that drives innovation and provides crucial knowledge about such pressing societal challenges as the effects of global change, biodiversity loss, sustainable food production, ecosystem conservation, and improving human health and security. Hormones can be extracted from decades-old natural history collections, making it possible to infer the physiological state of the individuals at the time of capture. Investigations using U.S. and international museum collections and private collections were the first to demonstrate how species respond to climate change by shifting locations, adapting to new conditions, or experiencing local extirpation. As new technologies and methodologies in research provide new insights about these specimens, sometimes making possible scientific uses never thought possible, the value of biological collections increases even more.

Biological collections are powerful educational assets for learners of all ages, backgrounds, skills, and perspectives. They provide a tangible platform that can draw people into *lifelong learning*—ongoing efforts to foster, develop, and expand one's knowledge and skills—whether through formal education, employment in science, technology, engineering, and mathematics (STEM), or by pursuing personal interests throughout life. By facilitating learning across a wide range of disciplines in formal and informal environments, biological collections can deepen subject-matter expertise and stimulate integrative and generative thinking that can link disciplines from the sciences to humanities and the arts. Biological collections can also inspire awe and stimulate curiosity, thus triggering questions not just about biology of individual organisms and species diversity, but also about agriculture, energy, medicine, public health, and many other issues of critical importance to humanity. Educators also use biological collections to drive inquiry-based learning[2] in order to improve skills necessary throughout life such as critical thinking, management, data interpretation, and problem-solving. Inquiry-based learning and undergraduate research experiences, such as those provided by some biological collections, also improve student understanding of STEM concepts and may be important mechanisms to encourage diverse communities to pursue careers in STEM.

Biological collections can be incorporated into classroom and non-school settings or serve as a means to provide research experience. Educational kits, classroom visits, field trips, summer camps, online courses, tutorials, blogs, citizen science programs, and teacher workshops are a few of the educational tools and programs created by biological collections staff. Because biological collections are tangible, they can provide a natural entry point to biology and biodiversity for people who may have limited experiences in nature. They are also an exceptional resource for building data literacy at all levels of the data life cycle—finding, generating, curating, evaluating, and using data. For example, the Biodiversity Literacy in Undergraduate Education[3] project uses data derived from natural history specimens to integrate data literacy teaching into undergraduate biology curricula. Finally, biological collections empower people from all walks of life to connect to and learn about nature, thus building wonder and providing a source of inspiration and appreciation for the natural world.

VISION FOR THE NEXT DECADE

The significance of biological collections as research infrastructure continues to grow in ways that were unanticipated 20 or even 10 years ago. With strategic thinking and steady resource investments, biological collections could continue to be at the heart of scientific advances and education for the foreseeable future. Looking ahead, the committee developed a common vision for the biological collections community in the next decade:

> *To provide long-term support for collections-based scientific research, instill a culture of proper stewardship for and access to biological specimens, build and grow biological collections to better represent global biodiversity in space and time, promote access to biological collections as important educational resources for the general public, and encourage the exchange of biological resources and knowledge.*

With this vision, the major aim of this report is to stimulate a national discussion regarding the goals and strategies needed to ensure that U.S. biological collections not only thrive but continue to grow throughout the 21st century and beyond. This expansive endeavor requires creative leadership that encompasses a wide range of perspectives and expertise to identify the needs of collections infrastructure and ensure their sustainability and growth.

---

[2] Inquiry-based learning is a student-centered learning and teaching approach in which students' questions (inquiries) and ideas are prioritized.

[3] See https://www.biodiversityliteracy.com.

RECOMMENDATIONS FOR THE NEXT STEPS

In this report, the committee first explores the ways that biological collections have contributed to society by advancing scientific discovery and innovation, enriching education, connecting non-professional communities to nature and science, and preserving Earth's natural science heritage. Then, the committee addresses how the biological collections community is working toward a common vision in light of today's challenges, recognizing that the future success of the biological collections community—curators, collection managers, directors, and users of biological collections—depends on addressing four interrelated issues:

1. upgrading and maintaining the physical infrastructure and the growth of collections;
2. developing and maintaining the tools and processes needed to transform digital data into an easily accessible, integrated platform as cyberinfrastructure[4] increases in complexity;
3. recruiting, training, and supporting a diverse workforce of the future; and
4. ensuring long-term financial sustainability.

Realizing this vision will require enhanced communication and collaboration within the biological collections community and beyond as well as a renewed and expanded commitment to maintain the diversity of biological collections, help them grow, and promote their use in scientific research and education. Following are the specific recommendations.

Building and Maintaining a Robust Infrastructure

Infrastructure includes not only the physical space and equipment used to house and maintain the specimens in a collection but also their accompanying data and the procedures governing their care. It includes the technologies to produce digital data and the cyberinfrastructure to store, analyze, and aggregate data with those of other collections through online portals. Finally, biological collections infrastructure includes the trained staff, students, and volunteers who acquire, curate, manage, ensure the quality of, and coordinate the scientific and educational uses of biological collections.

Owing to the diversity of collection types, there is no one-size-fits-all list of physical infrastructure requirements. The assessment of infrastructure needs to take place at the level of individual collections. Biological collections would benefit from an individualized strategic plan to outline how day-to-day needs

[4] Cyberinfrastructure, a term first used by NSF, encompasses the computing systems, repositories, advanced instruments, software, high-performance networks, and people that enable and support data acquisition, storage, management, integration, mining, analysis, visualization, and distribution (adapted from Stewart et al., 2010; https://scholarworks.iu.edu/dspace/handle/2022/12967).

will be met, including issues related to preventive maintenance and quality control, and also how to develop or expand infrastructure to meet future scientific needs.

Biological collections infrastructure also needs to grow in order to keep up with the advance and evolution of scientific research itself. The urgency to continue collecting will require NSF and other funding institutions, as well as institutions whose mandates include collecting or generating new types of research specimens, to acknowledge and address growth as an important and necessary component of biological collections in the 21st century.

Recommendation 4-1: The leadership (directors, curators, and managers) of biological collections should assess and define the infrastructure needs of their individual facilities and develop comprehensive strategic plans in accordance with those needs and their strategic missions. The strategic plans should outline approaches to:

- continually address ongoing preventive maintenance and, in the case of living collections, quality control requirements; and
- improve and potentially build new infrastructure, both of which are particularly important if collections growth is a component of the strategic mission.

The strategic plan should be revisited every 3 to 5 years to ensure that it continues to meet the evolving needs of collections and their users.

Recommendation 4-2: Biological collections should take advantage of existing training opportunities and collaborative platforms at the national and international levels, such as those offered through the International Society for Biological and Environmental Repositories and the Organisation for Economic Co-operation and Development certification programs, especially as new aspects of the work evolve, such as regulations compliance, data management, and new techniques and materials for collections storage and documentation.

Recommendation 4-3: Professional societies, associations, and coordination networks should collaborate and combine efforts aimed at addressing community-level infrastructure needs of the nation's biological collections, including:

- develop a platform to pool and share resources such as strategic plans, best practices, and training opportunities so that these can serve as resources for the broader biological collections community;
- develop and implement strategies to adopt quality control programs to improve uniformity among living stock collections and ensure the

availability of high-quality biological resources that best fit the needs of the user;

- create a national biological collections registry to document the location, size, and holdings of the collections in the United States. The registry should be curated and updatable. In addition, proactive processes to identify collections should be established, ensuring that collections of all types are well represented in the registry; and
- use the national registry to conduct periodic community-wide assessments of needs to inform the development of both individual and community-level strategies to maintain and upgrade infrastructure.

Recommendation 4-4: The National Science Foundation (NSF) Directorate for Biological Sciences should continue to provide funding support for biological collections infrastructure and expand endeavors to coordinate support within and beyond the Directorate. Specifically, NSF should:

- support new and improved infrastructure to accommodate the pressing needs created by continued collections growth;
- require a specimen management plan for all research proposals that includes collecting or generating specimens that describes how the specimens and associated data will be accessioned into and permanently maintained in an established biological collection; and
- facilitate the creation and support of an independent consortium to develop collaborative platforms and mechanisms to pool and share resources for strategic planning, preventive maintenance, quality control and assurance, collections growth, establishing a national collections registry, and other community-level assets.

Generating, Integrating, and Accessing Digital Data

Throughout their history, biological collections and the physical specimens they contain have been explicitly linked to the physical location where they are housed. To access the specimens and their accompanying written collections, users had to travel to a collection or receive specimens through the mail. Producing specimen data in digital formats is a vital first step toward enhancing the discoverability and use of biological collections. Digitization[5] and the cyberinfrastructure that underlies how digital data are stored, managed, and used have fundamentally transformed the biological collections community.

A key component of digitization has been the development of collection databases that provide digital specimen data to aggregated data repositories. Online data repositories facilitate the potential for new avenues of scientific inquiry, promoting the multiplication and expansion of research collaborations and community networks and providing a greater range of educational and training opportunities. A robust cyberinfrastructure can also facilitate evaluation and the development of metrics to assess the diversity of biological collections and their impact on research and education.

[5] The conversion of textual, image, or sound-based specimen information to digital formats.

Although digitization efforts have involved hundreds of collections, gaps in phylogenetic, geographic, temporal, and taxonomic information are evident. Investment in the development of new technologies and cost-effective high-throughput workflows for digitizing collections that, to date, have lagged—such as entomological collections—will enhance both the number of specimens and taxonomic scope of digitized collections.

A unified cyberinfrastructure that connects all types of biological collections, such as living and natural history collections, could accelerate research and provide innovative educational opportunities. Moreover, a permanent national cyberinfrastructure that supports the needs noted above in terms of expanded digitization of dark data, improvement in data quality, and increased accessibility to digital data would certainly spur data use. Without this resource, collections—both physical and digital—will continue to be underused.

The types of data that can be collected and their potential uses are beyond current imagination in terms of size, quality, complexity, and value. The "extended specimen" concept (see Figure S-2) opens the way to more opportunities, but implementing this concept requires both connecting with the research that uses the specimens and surmounting both technical and sociological issues of enabling and maintaining the linkage and inclusivity of the extended information through digital connections.

Recommendation 5-1: The leadership (directors, curators, and managers) of biological collections should provide the necessary mechanisms for staff to keep pace with advances in digitization and data management through training in digitization techniques and publishing of standardized quality data that can be efficiently integrated into portals.

Recommendation 5-2: Professional societies should initiate and cultivate opportunities for research collaborations within the biological collections community. These collaborations should include working with the computer and data science communities to promote the development and implementation of tools to build the cyberinfrastructure (e.g., data storage, annotation, integration, and accessibility to expand the use of biological collections to a broader range of stakeholders).

Recommendation 5-3: The National Science Foundation (NSF) Directorate for Biological Sciences should continue to provide funding for the digitization of biological collections and for the cyberinfrastructure to support both living and natural history collections. Specifically, the NSF Directorate for Biological Sciences should:

- partner with other directorates within NSF (e.g., physics, chemistry, computer science, and education) and other federal agencies and departments (e.g., the Department of Health and Human Services, the Department of Agriculture, the Food and Drug Administration, the Department of the Interior, the National Oceanic and Atmospheric

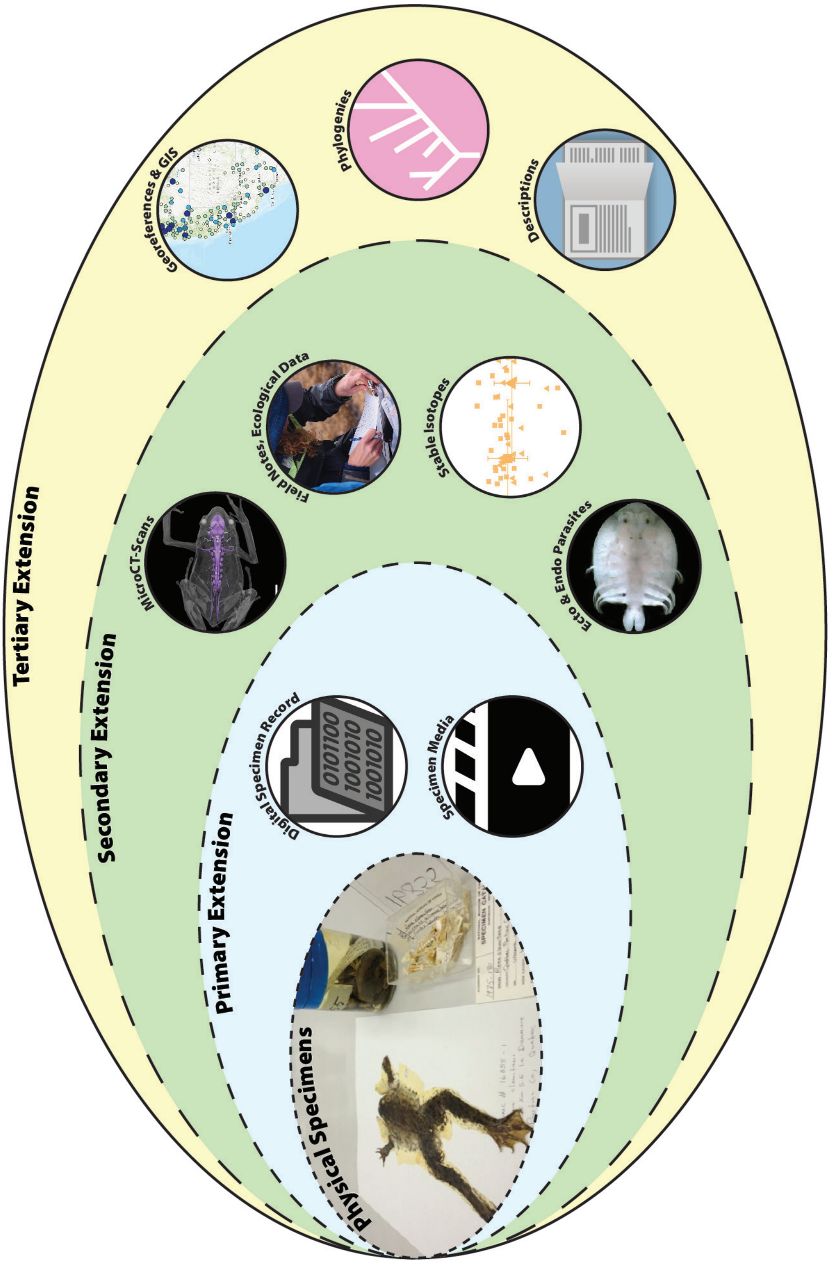

Tertiary Extension
Georeferences & GIS
Phylogenies
Descriptions
Secondary Extension
Field Notes, Ecological Data
Stable Isotopes
MicroCT-Scans
Ecto & Endo Parasites
Primary Extension
Digital Specimen Record
Specimen Media
Physical Specimens

FIGURE S-2 The Extended Specimen Concept. Extended specimens are collected and preserved in ways that encourage the use of different sets of analyses and questions. As detailed by Thiers et al. (2019), the extended specimen concept includes four components that in concert enable scientists to "capitalize on the depth and breadth of biodiversity held and digitally accessible in U.S. collections": (1) the physical specimen; (2) a primary extension that includes a digital record that brings together specimen-associated genotypic, phenotypic, and environmental data, including various media (e.g., images, sounds, video recordings); (3) a secondary extension that includes specimen-associated data that may be held in repositories or collections that are physically and digitally distinct and disconnected from the physical specimen such as isotype samples, gene sequences, or parasites found on the specimen; and (4) a third extension that includes data from other sources that may link to the physical specimen, such as descriptions and distribution of the species. Images of physical specimen (frog) courtesy of Dr. Kamal Khidas, Canadian Museum of Nature, Ottawa, Canada; digital specimen record icon by Jing.fm; specimen media icon by Gregor Cresnar, Flaticon.com; Micro-CT scan courtesy of David C. Blackburn and Edward L. Stanley, Florida Museum of Natural History; field notes picture by Mary Lewandowski; ecto and endo parasites image; and the georeferences map from the U.S. Geological Survey.

Administration, the National Aeronautics and Space Administration, the Department of Energy, etc.);
- establish ongoing mechanisms for the biological collections community to meet, develop best practices, and work toward goals such as establishing and implementing unique identifiers, clear workflows, and standardized data pipelines; and
- promote and fund the development of a necessary national cyberinfrastructure, with appropriate tools and technology to affect the efficient multi-layer integration of data and collections attribution.

Cultivating a Highly Skilled Workforce

If biological collections are to not just survive but thrive throughout the 21st century, they will need effective, visionary, and well-supported leaders, in addition to competent and innovative scientists and educators. Biological collections require personnel with multifaceted and complex competencies. Cultivating a highly skilled collections workforce, one that serves the data-intensive, globally connected, and often fast-paced needs of science and society, is essential to the long-term sustainability of the nation's biological collections.

The challenges facing biological collections are beyond the capability of any one institution to adequately address alone. A deeper understanding of the scope and needs of the existing collections workforce, identifying critical skillsets shared among the nation's biological collections, and building a sufficient workforce pipeline require collaborative, coordinated action. The path forward will require collaboration among the nation's biological collections as well as partnerships with other professional communities, incentivized with the support of NSF.

Recommendation 6-1: The leadership of individual collections, host institutions, relevant professional societies, and collections funders should collaborate to develop and strengthen the workforce pipeline through community-level action on the following issues:

- *Critical Skills*. Define critical, broadly applicable skillsets needed to lead, manage, and care for biological collections and expand and promote their uses for the national and global scientific enterprise and the benefit of society.
- *Workforce Analysis*. Conduct a comprehensive analysis of the existing collections workforce that, at a minimum, examines the professional responsibilities, demographics, education and training, incentives, compensation and benefits, and perceptions of greatest needs and opportunities for career development. Such an analysis should be conducted on a periodic basis (e.g., every 5 to 7 years) to inform community-level conversations and strategic action plans.
- *Diversity, Equity, and Inclusion*. Develop and implement programs to build a more diverse, equitable, and inclusive workforce. These programs

should include elements such as restructured classroom and mentoring practices, student internships, research opportunities to ensure that opportunities are more visible and accessible to diverse students and early-career professionals, and dedicated funding programs for internships and conference travel, workshops, and mentoring programs for diverse students and early-career professionals.

- *Education and Training Coherence*. Harmonize the design and offerings of biological collections–focused curricula and certificate and degree programs to fill current and future workforce education and training needs. This effort should include developing partnerships and cooperative arrangements with professional societies (e.g., for collections management training and taxonomic expertise), professional networks (e.g., in formal and informal education), and professional programs (e.g., museum studies, library studies, data science), respectively, to facilitate the design and implementation of biological collections–focused education and training programs in skillset areas not traditionally part of scientific training, and creating an online registry or portal to facilitate centralized access to information sharing about available education and professional development opportunities.
- *Alternative Staffing Models*. Provide guidance on alternative, innovative staffing strategies, including mechanisms to formalize student or volunteer involvement in collections management, that can help address staffing shortages, meet critical skillset needs, and serve as a mechanism to deepen collections knowledge among a broader range of people.

Recommendation 6-2: As part of its programmatic endeavors to promote a robust biological infrastructure, the National Science Foundation Directorate for Biological Sciences should support initiatives that focus explicitly on systemic, systematic, and thoughtful development of the biological collections workforce pipeline. In partnership with other directorates, such a programmatic focus should encompass future (e.g., students and postdocs) and existing collections personnel (e.g., early-career and senior curators and collection managers), and be predicated on maintenance and growth of biological collections infrastructure to meet diverse needs of societal import.

Securing Financial Sustainability

Long-term financial viability is critical to the ongoing and growing use of biological collections for research and innovation. Maintenance and replacement of aging physical infrastructure, continual upgrades to cyberinfrastructure, additional personnel to manage growing digital resources, upgrades to meet the needs of new emerging types of collections, new quality standards, and evolving requirements for permits and safety regulations are some of the funding needs that, while essential, may go beyond what annual budgets have covered historically.

Central to this effort is the development of comprehensive business plans that include estimates of the public funds needed to support the research that generated the collection and the infrastructure needs of the scientists that use collections as well as maintaining and providing access to the collections.

The biological collections community will need to act as one in order to develop partnerships, centralize a pooled set of data and resources, track the use of collections in research and education using diverse metrics at the community level to show the national and international impact of U.S. collections, and identify new approaches to funding.

Recommendation 7-1: The leadership (directors, curators, and managers) of biological collections should work with business strategists and communication experts to develop business models for financial sustainability and infrastructure of biological collections. Included in this discussion should be the development of a mechanism to

- diversify funding portfolios and develop relationships with non-traditional partners who may provide collections support;
- assess a per-specimen acquisition and maintenance cost. This assessment would depend on the size and nature of the collection—both physical and digital; and
- explore revenue streams that could include pay-for-use models, the establishment of material transfer agreements and licensing systems, or perhaps pay for value-added for digital datasets configured for a particular purpose. Each of these approaches must be done in ways that avoid driving costs to levels that are prohibitive for researchers.

Recommendation 7-2: Professional societies should develop extensive networked training platforms for sharing best practices for financial management and planning and business models for collections of all sizes and types. This could be an ongoing activity centered at a national biological collections center and should include both natural history and living collections together.

Recommendation 7-3: The National Science Foundation Directorate for Biological Sciences should continue to provide stable, long-term funding to support investigators who rely on biological collections for research and education. Specifically, it should:

- work with other federal agencies to address research infrastructure support and needs;
- provide funding for the management and infrastructure of the collections themselves;
- collaborate with host institutions and other funders to establish new mechanisms and funding to collect, aggregate, and synthesize metrics to evaluate process and performance for biological collections; and

- support the accessioning, curation, digitization, and long-term care of specimens as well as the publishing of their associated data through a mandated specimen management plan.

Taking Collaborative Action

There is a growing recognition that integrated global initiatives that leverage diverse perspectives, institutions, and resources are needed to prevent and respond to issues of high international priority such as emerging infectious diseases, biodiversity loss, food security, invasive species, or climate change. If more fully connected across diverse disciplines, biological collections could play a much larger role in these initiatives.

Coordination and sharing of knowledge will be critical for the biological collections community to be able to meet current and future needs and address the dynamic challenges of society and rapid global change. The biological collections community needs an inclusive, integrated platform to strengthen the position of biological collections as a unified scientific infrastructure for the nation over the next decade and beyond. A national collections-focused action center dedicated to the support and use of biological collections could fill this need.

Recommendation 8-1: The National Science Foundation, in collaboration with other institutions that provide funding and other types of support for biological collections, should help establish a permanent national Action Center for Biological Collections to coordinate action and knowledge, resources, and data sharing among the nation's biological collections as they strive to meet the complex and often unpredictable needs of science and society. Such an action center should include a physical space and cyberinfrastructure to develop and implement collaborative strategic efforts and further build and nurture communities of practice for research, education, workforce training, evaluation, and business model development, among other community-wide needs.

Recommendation 8-2: The National Science Foundation should lead efforts to develop a vision and strategy, such as a decadal survey, for targeted growth of the nation's biological collections, their infrastructure, and their ability to serve a broader range of users and scientific and educational needs. The vision and strategy should take into consideration the diverse capabilities and needs of all types of collections and diverse array of end users, and set long-range priorities that could only be accomplished with a concerted, collaborative effort of the nation's biological collections.

Recommendation 8-3: The National Science Foundation (NSF) Directorate for Biological Sciences should expand its partnership capabilities more broadly across NSF, other federal agencies, international programs, and other sectors. Such partnerships can maximize investments in support of a national Action Center for Biological Collections and the development of a national vision and strategy and help spread the cost of such major endeavors beyond the NSF Directorate for Biological Sciences.

1
The Repository of Life

Life comes in many forms, sizes, and shapes. This rich diversity of forms, sizes, and shapes of life on Earth, estimated at more than 1 trillion species (Locey et al., 2016), gives rise to wonder and fuels the curiosity that drives scientific discovery, advances, and innovation worldwide. For centuries, scientists have sought and collected different types of organisms to learn more about their forms, functions, origins, distributions, and evolution. Pooling and conserving these organisms into *biological collections*—systematized repositories of life in all of its many forms—is a cornerstone of quality research and education in many areas of science and innovation (Dunnum et al., 2017; Jarrett and McCluskey, 2019; Koornneef and Meinke, 2010; McCluskey, 2017; Meineke et al., 2018b; Schindel and Cook, 2018). Scientists and educators who study and teach about life on Earth rely on biological collections as an important underlying scientific infrastructure upon which their knowledge and learning build and grow.

Biological collections typically consist of organisms (specimens) and their associated biological material, such as preserved tissue and DNA, along with data—digital and analog (such as handwritten field notes)—that are linked to each specimen. Non-living specimens include organisms preserved by scientists and naturally preserved remains, such as fossils. Such collections of non-living specimens are commonly referred to as natural history collections. Living specimens include research and model organisms that are grown and maintained in genetic stock centers, germplasm repositories, or living biodiversity collections. The defining trait of these different types of collections is that they capture aspects of the living world in such a way that it can be intensively studied and understood through time.

Biological collections provide a wide range of benefits to science and society. For one, biological collections are at the core of dynamic research on globally relevant societal issues by serving as archives of our natural heritage and preventing loss of knowledge about life on Earth. They support research on basic biological structures and processes (e.g., Lister, 2011; Shaffer et al., 1998) and deepen our understanding of evolution, biodiversity, and global environmental change (Lang et al., 2019; Meineke et al., 2018b). Herbarium[1] specimens, for example, can be used to study atmospheric conditions in the past and inform scientific understanding of global change over time (see Box 1-1). Biological collections advance science in ways unanticipated from when a specimen was first collected. One renowned example is the development of the polymerase chain reaction technique for replicating DNA, which was among the most influential discoveries of the 20th century (see Box 1-2). Biological collections also underpin and enrich the knowledge of students of all ages about biology and biodiversity (Antunes et

[1] Natural history collections of plants.

BOX 1-1
Stomata: Hints of Atmospheric Conditions in Past Times

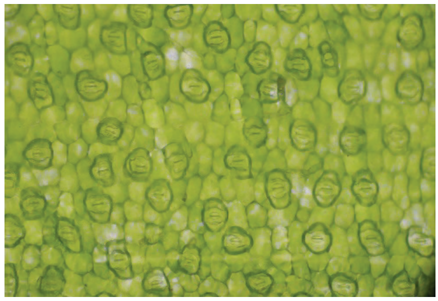


Image by toeytoey2530 on iStockphoto.com

A natural history collection can serve as a "snapshot" of biodiversity at the time that the collection was made. Multiple collections of similar material over a long timeframe create a veritable photo album that can chart important ecosystem changes over past decades, centuries, or millennia. For example, plants contain structures that can tell scientists about historical atmospheric conditions. These structures, known as stomata, are small holes on the underside of leaves that permit the exchange of gases, including carbon dioxide, associated with the process of photosynthesis.

F. Ian Woodward hypothesized that the higher the atmospheric carbon dioxide, the lower the number of stomata. He conducted controlled experiments to demonstrate the effect of atmospheric carbon dioxide on the density of stomata on plant leaves. As carbon dioxide levels increased, fewer stomata were needed for gas exchange to fuel photosynthesis.

Earth's history since life began is characterized by wide swings in atmospheric carbon dioxide, and scientists note that increasing levels of carbon dioxide from human activities are leading to global warming and other changes. If one could examine leaves from 100 or 200 years ago, scientists reason, the relative abundance of stomata would be a good proxy for how much carbon dioxide was in the atmosphere when the plant was alive. Woodward turned to herbarium specimens held by the Department of Botany at the University of Cambridge. Using selected tree species, he examined the density of stomata over the past several hundred years. He found that the average density of the stomata had dropped by 40 percent over the past 200 years, adapting to the increased availability of atmospheric carbon dioxide (Woodward, 1987).

Given the long record of plants on Earth, though, it would be desirable to be able to go back before herbarium collections existed. Fortunately, collections can help there, too—fossil collections. Later studies expanded on Woodward's work and continued with a project to examine stomatal density in fossil plants, which provided evidence about how changes in atmospheric carbon dioxide levels may have affected biodiversity in prehistoric times (Soul et al., 2018).

BOX 1-2
***Thermus aquaticus*: Breaking Biological Barriers**

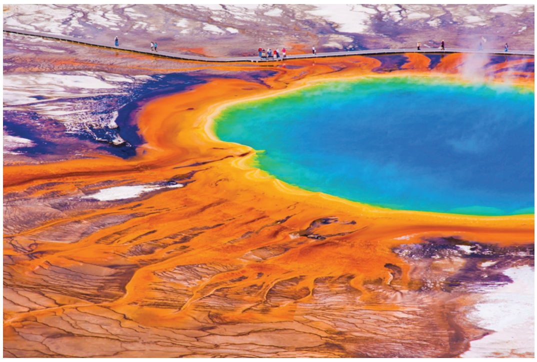


Image by lorcel on iStockphoto.com

In the mid-1960s, Thomas Brock, a microbiologist, and his undergraduate student Hudson Freeze made an unanticipated discovery. With support from the National Science Foundation, Brock was collecting and studying heat-loving microorganisms from hot springs and geysers in Yellowstone National Park. He was interested in the influence of extreme heat on photosynthesis and primary production in cyanobacteria (Brock, 1967; Brock and Brock, 1967, 1968). During this time, it was believed that bacteria could not live at temperatures above 55ºC and that the upper temperature threshold for life in general was 73ºC. However, Brock discovered microorganisms thriving at temperatures hotter than ever known to be possible. In 1966, Brock and Freeze isolated and cultured one of these heat-loving bacteria, which they named *Thermus aquaticus* (Brock and Freeze, 1969; Freeze and Brock, 1970). Later, Brock and another student, Gregory Zeikus, demonstrated that enzymes from this microbe could also tolerate extremely high temperatures (Zeikus and Brock, 1971). Unbeknownst to Brock and his students, in addition to upending assumptions about the conditions in which life could thrive, their findings would also become the bedrock of modern biotechnology.

More than 20 years after the discovery of *T. aquaticus*, another scientist, Kerry Mullis, had a great idea but could not figure out how to make it work at a larger scale. In the early 1980s, Mullis, a biochemist, wanted a fast and efficient way to make copies—lots of them—of specific bits of DNA, often bits where the original sample was very small. "Amplifying" genes of interest would give scientists a way to identify, study, or manipulate them, and to share them in quantity with colleagues. Mullis was working with *Escherichia coli*, but the process in use at the time required cycle after cycle of heating to break apart the original DNA strands to amplify them, and every time he heated the *E. coli*, its DNA polymerase fell apart.

Mullis realized that he needed a particular type of enzyme that could survive and function at high temperatures (Mullis, 1990). He knew that some bacteria

continued

BOX 1-2 Continued

could withstand much higher temperatures than *E. coli*: The problem was how to find some without going to a thermal vent or a hot spring and hunting for a suitable microbe. Fortunately, Brock and Freeze had already done the legwork, and crucially for Mullis, Brock had also sent live samples of the heat-loving *T. aquaticus* to the American Type Culture Collection (Innis et al., 1988), where it is still housed today.

The technique Mullis perfected using *T. aquaticus* from a living biological collection—which we now know as the polymerase chain reaction—is the foundation of modern biotechnology and biomedicine, used in routine lab tests in doctors' offices, in DNA fingerprinting to solve crimes, in re-creating DNA from extinct plants and animals, and in performing rapid diagnoses of infectious diseases. The Nobel Prize for this breakthrough went to Mullis in 1993.

al., 2016; Beckmann et al., 2015; Lacey et al., 2017). Schools, universities, and research laboratories use biological collections to teach concepts of evolution, ecology, taxonomy, physiology, biogeography, conservation, and more (see Box 1-3). Finally, many biological collections connect the public to nature and science, bolstering lifelong learning (Graham et al., 2004; Hill et al., 2012; MacFadden, 2019; Suarez and Tsutsui, 2004).

Unfortunately, the sustainability of the nation's biological collections is under threat. The causes are many, ranging from a general lack of understanding of their value and their contributions to research and education and a lack of appreciation for what is required to maintain them effectively, to inadequate coordination and interconnection among the collections that make up the critical infrastructure. Without necessary changes in support and leadership, the prior and current investments in time, money, and staff resources for building the nation's biological collections will be diminished, and their immense potential in supporting science, innovation, and education in the United States and elsewhere will be severely limited.

PURPOSE OF THE STUDY

Recognizing the importance and the vulnerabilities of the nation's biological collections, the National Science Foundation (NSF) has endeavored to provide broad financial support through its Division of Biological Infrastructure (DBI) within the Directorate for Biological Sciences (BIO) (see Box 1-4). However, the breadth of needs for maintaining biological collections exceeds the capabilities of any one federal agency. Many U.S. government agencies, including NSF, the Department of Agriculture (USDA), the National Institutes of Health, the Department of the Interior, and the Department of Energy, support research that uses and creates biological collections, but agency support for maintenance

BOX 1-3
Biological Collections as Educational Resources: The Marine Resources Center

Top: Image by Dodds, S. Gideon; Bottom: Image by Megan Costello

Biological collections are a powerful resource for both formal and informal education. At many U.S. universities, natural history collections expose students to the diversity of life and form the foundation for teaching concepts of evolution, ecology, taxonomy, and more. But biological collections have an even greater reach through museums, field stations, and research laboratories where learners of all ages can explore specimens both physically and virtually. Some research facilities house living collections, with unique opportunities for research training in a host of basic and applied disciplines.

The Marine Biological Laboratory (MBL; https://www.mbl.edu) at Woods Hole, Massachusetts, has offered formal and informal educational programs since 1888. Today these programs are supported by the living collections of the Marine Resources Center (MRC; https://www.mbl.edu/mrc), which maintains, cultures, and provides aquatic and marine organisms for both research and

continued

BOX 1-3 Continued

education. Although a key source of research materials for science laboratories worldwide, the MRC collections of fish, frogs, mollusks, and more play a complementary role in formal and informal education.

Courses at the MRC use both living stocks of model organisms such as zebrafish or frogs and locally collected samples of marine life to provide interdisciplinary research training for students from high school through graduate school. Field courses—from summer camps to tours to university programs—introduce students of all ages to marine biodiversity through collecting, observation, and hands-on research. Together, these formal and informal activities at MBL, as at other institutions nationwide, can be important catalysts for attracting students to careers in science (Elkins and Elkins, 2007; Pawson and Teather, 2002).

BOX 1-4
The National Science Foundation Support for Biological Collections Infrastructure

For many decades, the National Science Foundation (NSF) has been a vital source of support for biological collections. Currently, the NSF Directorate for Biological Sciences (BIO) has two ongoing support programs: (1) Collections in Support of Biological Research (CSBR) and (2) Advancing Digitization of Biodiversity Collections (ADBC). The goal of the CSBR program is to strengthen the infrastructure essential to carrying out research in the areas of interest to NSF/BIO—the principles and mechanisms governing life across all scales of biological organization, from molecules to ecosystems to the global biosphere. CSBR provides funds for three general infrastructure needs: (1) improvements to secure and organize collections that are significant to the NSF/BIO-funded research community; (2) securing collections-related data for sustained, accurate, and efficient accessibility to the biological research community; and (3) transferring ownership of collections. ADBC provides support for expanding and enhancing digital natural history collections data and improving access to digitized information. The NSF Directorate of Earth Sciences also contributes funding to CSBR when there is a relevant proposal and pending available funding. Until 2011, infrastructure for living collections and natural history collections were supported through separate solicitations. In 2017, NSF suspended its collections infrastructure support program, which sparked an outcry from the scientific community (Nowogrodzki, 2016a; Rogers, 2016) and led to the infrastructure support program being merged and reinstated in 2018, but at a lower funding level than it had been earlier (Nowogrodzki, 2016b). In 2019, NSF/BIO initiated a program, Sustained Availability of Biological Infrastructure, that includes provisions for supporting biological living stocks that face ongoing operational costs that exceed those available from their host institutions.

of those collections, if any, is not proportional to their use in agency-funded research. NSF is continuing to provide support, but it welcomes guidance on a wide range of questions. What operational structures, policies, and cultures could provide momentum to maintain and grow biological collections? What options are adaptable, transferable, or scalable for different types of collections? What is needed to ensure the long-term sustainability of the nation's biological collections? For these reasons, NSF asked the National Academies of Sciences, Engineering, and Medicine (the National Academies) to address the following:

- explore the contributions of biological collections of all sizes and institutional types to research and education;
- envision future innovative ways in which biological collections can be used to further advance science;
- outline the critical challenges to and needs for their use and maintenance, including the quality control challenges faced by living stock collections, to enable their continued use to benefit science and society; and
- suggest a range of long-term strategies that could be used for their sustained support.

The full Statement of Task for the study is provided in Appendix A. NSF asked that these tasks be addressed in the context of the "living stocks (organisms) and preserved repositories of biodiversity specimens and materials" (i.e., natural history collections) that receive, or are eligible to receive, support for infrastructure or digitization from NSF-DBI. As a result, this report does not *explicitly* address living collections in zoos, aquaria, or botanical gardens; biobanks or repositories of human tissues; or anthropological and geological collections (excluding fossils). This report does not cover biological collections owned by federal agencies. Although these types of collections may be housed in the same institutions as NSF-supported biological collections or be used in research supported by NSF (e.g., USDA germplasm collections), DBI does not provide support for their infrastructure. The committee, however, recognizes that many of the "excluded" collections share the same challenges and opportunities. Thus, examples used in the report may be drawn from collections outside the domain of NSF-DBI-supported research.

The Committee's Approach

To fulfill the Statement of Task, the National Academies convened a committee of 13 distinguished experts whose collective experience included a diversity of biological collections, K–12 and informal education, and science communication. The committee held four in-person meetings, including a public workshop, and five webinars as part of its information-gathering process (see Appendix B for the public meeting agendas and list of invited speakers). The public meetings, workshop, and webinars featured a total of 25 speakers who covered a range of topics needed to address the Statement of Task, including the history, philosophy, and role of biological collections; emerging and novel applications of biological collections in research and education; and advances in cyberinfrastructure and digitization. As befits an issue of great concern to the Earth and life sciences communities, a number of experts have issued reports describing the challenges facing both federal and non-federal collections in the United States and identifying opportunities for integration, innovation, and tracking long-term impacts (see Box 1-5). These reports address specific categories of biological collections: a total of six reports on federal biological collections, geological collections, living stock collections, genetic collections, and natural history collections. The committee's analyses and deliberations led to this final Consensus Study Report, which draws on the presentations the committee heard, its review of scientific and other literature, and the expertise of its members.

In responding to the Statement of Task, the committee considered two broad categories of biological collections: (1) non-living organisms, also referred to as natural history collections; and (2) living organisms, including research and model organisms. The scope of the study is broad, encompassing the contributions of "biological collections of all sizes and across institution types to research and education." The committee identified areas of tension that stem from the scope of the study and that are inherent within the biological collections community. Biological collections are diverse—taxonomically, organizationally, and in their missions and needs. There is also tension that arises from differences between living stock collections and natural history collections. With the exception of a few biodiversity-focused living collections,[2] living and natural history collections communities (e.g., directors, managers, curators, and users) operate largely independently of one another. This report is the first of its kind to address the challenges and promise of both living stock collections and natural history collections. The committee acknowledges that living stock collections and natural history collections have distinct purposes and needs, but the committee also found that there are many opportunities for these communities to learn from one another and collaborate. Throughout the report, the committee highlights some of these potential synergies and intersections (e.g., digital genetic data, extended specimen information) as well as key distinctions (e.g., business strategies, quality control). The report is not an exhaustive compendium of every issue, but is intended to launch a national conversation about strategic collaboration between the living stock and natural history collection communities.

THE PROMISE OF BIOLOGICAL COLLECTIONS

Biological collections are an invaluable, and often irreplaceable, component of the nation's scientific enterprise. They are a rich and diverse data source providing the research and education communities with keys to decoding the living world—past, present, and future. For hundreds of years biological collections have inspired and informed science, but their promise has never been greater than it is today. Part of that increase in scientific value can be attributed

[2] The Duke Lemur Center, Durham, North Carolina (fossil collections), or the Montgomery Botanical Center, Coral Gables, Florida (herbarium), are examples of living biodiversity collections that interact with in-house natural history collections.

BOX 1-5
Selected Reports on Importance and Needs of Biological Collections in the United States

The Biological Resources of Model Organisms (Jarrett and McCluskey, 2019)
This book provides a brief look at the individual organisms, how they came to be accepted as model organisms, the history of the individual collections, examples of how the organisms have been and are being used in scientific research, and a description of the facilities and procedures used to maintain them.

Extending U.S. Biodiversity Collections to Promote Research and Education (Thiers et al., 2019)
This report is the result of a consensus discussion, led by the Biodiversity Collections Network, on the future of biodiversity data held in U.S. biological collections. The report recommends building a network of extended specimens to facilitate research across taxonomic, temporal, and geospatial scales.

Scientific Collections: Mission-Critical Infrastructure for Federal Science Agencies (IWGSC, 2009)
This report focuses on U.S. federal object-based scientific collections, including biological collections. Written by the Interagency Working Group on Scientific Collections, this report describes the diversity and purpose of federal scientific collections and makes recommendations for ongoing responsible stewardship.

Geoscience Data and Collections: National Resources in Peril (NRC, 2002)
This Consensus Study Report of the National Research Council outlines a comprehensive strategy for managing geoscience data and collections (including fossils of all types) in the United States.

The U.S. National Plant Germplasm System (NRC, 1991) and ***Managing Global Genetic Resources: Agricultural Crop Issues and Policies*** (NRC, 1993b)
This Consensus Study Report series examines needs and approaches in preserving genetic material for agriculture, including the worldwide network of genetic collections, the role of biotechnology, and a host of issues that surround management and use.

simply to the steady growth in the collections over time, but other factors have played major roles in their value: the growing diversity of biological collections, the development of new technologies to study collections, and the explosion of digitization of collections over the past few decades.

Diversity of Biological Collections

Today's biological collections are highly diverse—they exist in distributed physical locations and vary in size, taxonomic diversity, origin, the kinds of specimens and data generated, and how they are maintained and used (see Figure 1-1). Typically, a collection consists of physical groupings of living or preserved organisms and selected and curated parts of those organisms, such as tissue, blood, or DNA (Ankeny, 2019), together with the comprehensive data associated with the specimens. Many institutions house biological collections from multiple taxonomic groups from around the world and across multiple geological timescales. Other biological collections consist of genetically modified microbes, plants, vertebrates, or invertebrates used for their diversity in genotypes, phenotypes, and physiological functions, regardless of where they originated. Variety in collections and how they are used is a recurring theme throughout this report. While this report covers only certain kinds of collections (see section on the scope of the report), collections can range in size from millions of specimens in large collections to smaller, project-based[3] collections. They are housed in natural history or science museums, botanical gardens, universities, biological resource or stock centers, or private or even small collections of the sort that result from the efforts of one or a few investigators working on a single project. The scientific literature is replete with research made possible only, or primarily, because of biological collections and their unique combination of biological material and associated data. Examples of specific ways in which biological collections contribute to research and education can be found throughout this report, with unique contributions highlighted in this chapter and in Boxes 2-1, 2-2, 3-1, 3-2, 4-1, 4-5, and 5-1.

This tremendous diversity is both the single greatest asset of collections and the single biggest challenge they face. There is no one-size-fits-all solution for the myriad kinds of, and management approaches to, biological collections. Even the term "biological collection" often eludes a succinct description. For this report, the committee focused on collections developed for research, although

[3] Project-based biological collections (sometimes called ad hoc collections) are those generated for a specific research study. They usually do not continue to grow once the research concludes, and they typically lack funding for long-term maintenance or dedicated facilities to house them if the principal investigator retires or moves to a new institution, leaving the collection behind. Depending on quality and funding, some project-based collections may be maintained by their host institutions for new research purposes or transferred to a more comprehensive long-term repository.

FIGURE 1-1 Examples of biological collections in the United States. (A) spider in amber, University of Colorado Museum of Natural History Paleontology Section; (B) bats, Museum of Southwestern Biology, The University of New Mexico; (C) *Fusarium graminearum*, Fungal Genetics Stock Center, Kansas State University; (D) Xenopus, The National Xenopus Resource, Marine Biological Laboratory; (E) various herbarium specimens, New York Botanical Garden C.V. Starr Virtual Herbarium; (F) Charles Doe egg collection, Florida Museum of Natural History; (G) Ichthyology Cleared and Stained specimens in jars, University of Kansas Biodiversity Institute & Natural History Museum; (H) bacterial strain on petri dishes, American Type Culture Collection.

many research collections are used for formal and informal science, technology, engineering, and mathematics (STEM) education.[4]

Digitization of Biological Collections

Digitization, or the conversion of specimen information to digital formats, including high-resolution images and genetic sequence data, has improved the value and usability of biological collections in a number of ways. For instance, it provides quick, easy, and inexpensive access to millions of specimens as well as to myriad associated data for any users with an Internet connection (Soltis, 2017). As observed with living stock collections, such as the microbe collections listed in the Global Catalogue of Microorganisms (GCM),[5] the surge of available digital information for natural history collections is resulting in an increase of users of these collections and will undoubtedly spur research innovations in all disciplines of science (see Chapter 5). The countless available databases linked to specimens extend the concept of biological collections and enable novel

[4] In this report the committee adopts the NSF definition of STEM, which includes mathematics, natural science, engineering, computer and information science, and the social and behavioral sciences—psychology, economics, sociology, and political science (NSF, 2018).

[5] See http://gcm.wfcc.info.

specimen-based and new data-driven lines of scientific inquiry. The accessibility of databases of biological information mobilizes both basic and applied research (Nelson and Ellis, 2018) and has led to Nobel Prize–winning discoveries (see Box 2-1 and McCluskey, 2017).

The digitization of biological collections has also revolutionized the ability to distribute and share information from these collections. For centuries, scientists wanting to study a particular specimen from a natural history collection had to visit the place where it was held or have the specimen sent to them, leaving the item susceptible to loss or damage (Olsen, 2015). Today, the coordinated worldwide efforts to digitize biological collections and associated data (e.g., Integrated Digitized Biocollections[6] (iDigBio) funded through the NSF Advancing Digitization of Biodiversity Collections[7] (ADBC) program, and the European Distributed System of Scientific Collections[8] and Innovation and Consolidation for Large Scale Digitisation of Natural Heritage[9] program) provide access to rich sources of site- and species-specific data through data aggregators (e.g., Global Biodiversity Information Facility,[10] iDigBio, and GCM), which fuel innovative thinking (see Chapter 5). The advent of advanced technologies and computerized methods augments the physical specimens in biological collections with a wealth of digitized data as well as derived resources and metadata, both physical and digital.

These new approaches to generating, storing, and sharing specimens and their associated data not only enable specimen-based research but also make possible new approaches to solving complex global problems. Researchers have spoken of this as the "holistic" (Cook et al., 2016) or "extended specimen" concept (Webster, 2017) (see Figure 1-2). For users of living collections, genetic stocks act as repositories and distributors of biological specimens and their derived genotypic and phenotypic data and serve as a central hub for wide-ranging research communities. A specimen's aggregated data can be combined "to form an information-rich network for exploring Earth's biota across taxonomic, temporal and spatial scales" as recently noted in a report from the Biodiversity Collections Network (Thiers et al., 2019, p. 2).

The types of data that can be collected and their potential uses are beyond current imagination in terms of size, quality, complexity, and value. The "extended specimen" concept opens the way to more opportunities, but implementing this concept requires both connecting with the research that uses the specimens and surmounting both technical and sociological issues of enabling and maintaining the linkage and inclusivity of the extended information through digital connections. Given the immense number of sources of digitized biological information from all kinds of biological collections, mechanisms to inventory and evaluate the capabilities of biological collections in the United States and abroad are needed. This is a daunting challenge in a historically siloed world. Garnering,

[6] See https://www.idigbio.org.
[7] See https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503559.
[8] See https://www.dissco.eu.
[9] See https://icedig.eu.
[10] See https://www.gbif.org.

organizing, and aggregating this essential information is key to realizing a digital revolution. Harnessing the expansion of digital tools and technologies—online through accessible databases—empowers researchers to forge new links and open new avenues of inquiry, broadens education opportunities at all levels, and gives us the tools to embrace globalization.

The Value of Today's—and Tomorrow's—Biological Collections

The wealth and diversity of biological collections and their extended networks make it possible to approach issues of global importance holistically, bridging cultural and knowledge gaps. But biological collections also have catalyzed scientific discovery across a wide variety of fields, from medicine and public health to agriculture, ecology, evolutionary biology, and global change. For example, genetic stock collections of plants, insects, and microorganisms played a central role in advances in the field of genetics and applications to plant and animal agriculture (NRC, 1993a).

Biological collections provide a fundamental underpinning for a tremendous amount of basic research in the biological sciences (see Chapter 2). Consider, for instance, the revolutionary genome-editing technique known as CRISPR (clustered regularly interspaced short palindromic repeats). CRISPR has vastly expanded the genetic resources available in living collections and advanced the applications of biotechnology in medicine, agriculture, and conservation. Furthermore, the development of CRISPR was in part the result of research on materials sourced from living microbe collections (Ishino et al., 1987; Jinek et al., 2012). More generally, decades of groundbreaking life science research were only made possible because of the availability of high-quality living stocks and model organisms (McKie, 2017; see also Box 2-1).

Biological collections also help scientists predict and respond to a rapidly changing world. They have the unique capacity to validate existing research endeavors, reveal large-scale temporal patterns, and allow the retracing of environmental disturbances over time. For example, recent important insights into the effects of climate change on the distribution of mountain and desert organisms have been the result of comparisons of biological collections sampled and compared across a century of environmental change (Grinnell Resurvey Project;[11] Shaffer et al., 1998). Sometimes the connection between the biological collection and an outcome is reasonably straightforward, as when paleontologists study the fossils in a collection to gain insight into the evolution of a species or biologists use historical collections of plants or animals to understand how the geographic distribution of a species has changed over time. A recent example of the latter was research on the endangered Poweshiek skipperling (see Box 1-6).

The ability to collect vital, invaluable clues on disease patterns in humans, animals, and crops also depends on well-documented archived or reference biological collections (Ristaino, 2002). In many cases, analyses of both living stock and natural history collections are essential for public health officials to

[11] See http://mvz.berkeley.edu/Grinnell/pubs.html.

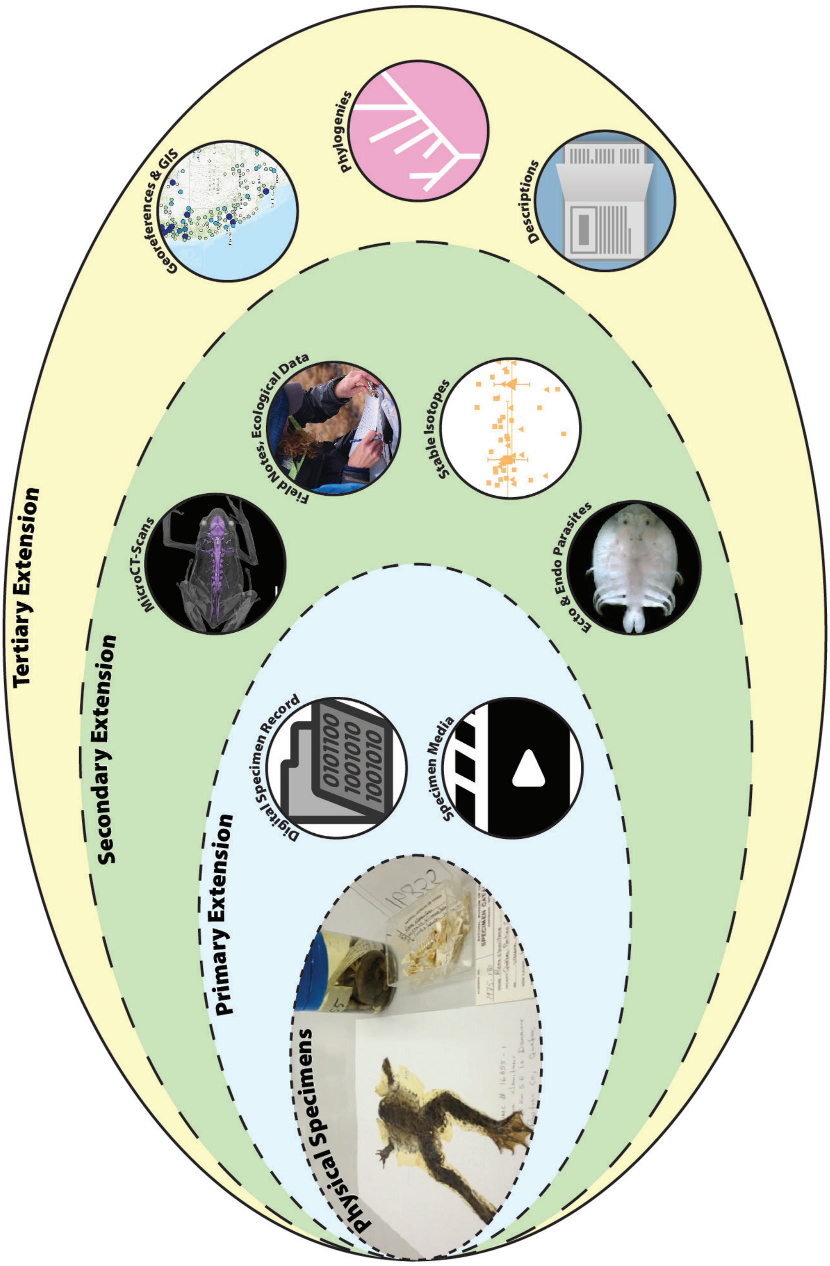

Phylogenies
Georeferences & GIS
Descriptions
Field Notes, Ecological Data
Stable Isotopes
MicroCT-Scans
Ecto & Endo Parasites
Tertiary Extension
Secondary Extension
Digital Specimen Record
Specimen Media
Primary Extension
Physical Specimens

FIGURE 1-2 The Extended Specimen Concept. Extended specimens are collected and preserved in ways that encourage the use of different sets of analyses and questions. As detailed by Thiers et al. (2019), the extended specimen concept includes four components that in concert enable scientists to "capitalize on the depth and breadth of biodiversity held and digitally accessible in U.S. collections": (1) the physical specimen; (2) a primary extension that includes a digital record that brings together specimen-associated genotypic, phenotypic, and environmental data, including various media (e.g., images, sounds, video recordings); (3) a secondary extension that includes specimen-associated data that may be held in repositories or collections that are physically and digitally distinct and disconnected from the physical specimen such as isotype samples, gene sequences, or parasites found on the specimen; and (4) a third extension that includes data from other sources that may link to the physical specimen, such as descriptions and distribution of the species. Images of physical specimen (frog) courtesy of Dr. Kamal Khidas, Canadian Museum of Nature, Ottawa, Canada; digital specimen record icon by Jing.fm; specimen media icon by Gregor Cresnar, Flaticon.com; MicroCT-scan courtesy of David C. Blackburn and Edward L. Stanley, Florida Museum of Natural History; field notes picture by Mary Lewandowski; ecto and endo parasites image; and the georeferences map from the U.S. Geological Survey.

BOX 1-6
Building a Database from Scratch: Poweshiek Skipperling

Image courtesy of Vince Cavalieri, U.S. Fish & Wildlife Service

The Poweshiek skipperling is an orange-brown prairie butterfly not much larger than a quarter, whose population crashed between 2005 and 2015. Saving the butterfly required learning more about its ecological niche, so a group of naturalists and ecologists set out to map its presence over time, from the second half of the 19th century to modern times—an effort that required poring through dozens of natural history collections and records. Today ecologists are using the assembled data to develop plans for bringing the Poweshiek skipperling back from the brink of extinction (Belitz et al., 2018; Pogue et al., 2016).

identify emerging pathogens and develop preparedness strategies to mitigate the spread of disease around the world (Shrivastava et al., 2018; Yanagihara et al., 2014). This report was produced in the middle of the coronavirus disease 2019 (COVID-19) global pandemic, which provides a timely example of how living and natural history collections infrastructure can be integrated to detect, describe, and mitigate emerging infectious diseases. During outbreaks and pandemics, living stock collections, such as the American Type Culture Collection and some of the government contracts they manage,[12] maintain and distribute virus strains and associated materials for basic research and development of diagnostic tests, therapeutics, vaccines, and detection methods. What is less obvious is the value of continuously using natural history collections infrastructure to better understand pathogen emergence on a global scale (Cook et al., 2020; DiEuliis et al., 2016;

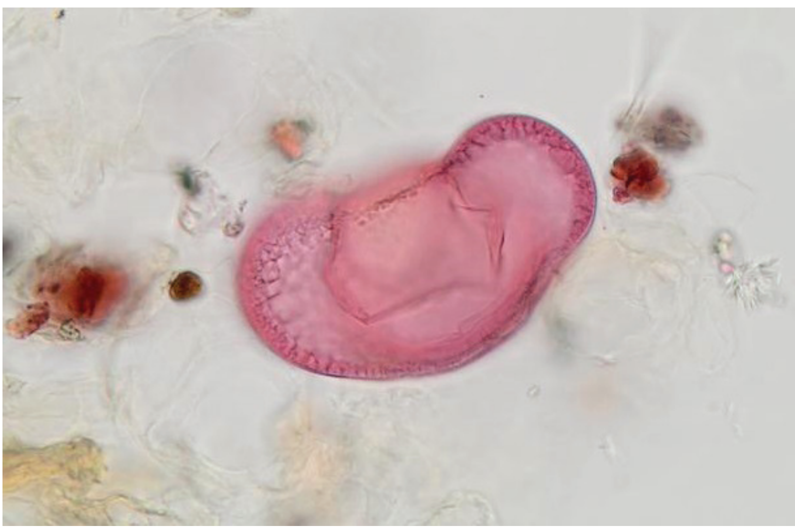
[12] The Biodefense and Emerging Infectious Resources has been funded in whole or in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN272201600013C and the International Reagent Resource has been funded by the Centers for Disease Control and Prevention.

Dunnum et al., 2017). Natural history collections are an essential resource for studying pathogen hosts and their spatial and temporal distribution (Harmon et al., 2019).

Many applications of biological collections also rely on making connections that are less than obvious, such as the use of pollen collections to help identify "Baby Doe," a young girl whose body was found in a plastic bag washed up on a Massachusetts shore (see Box 1-7).

Biological collections can inspire wonder, curiosity, and connectivity to nature in young and old, scientists and non-scientists alike, through formal

BOX 1-7
Pollen Forensics: Identification of "Baby Doe"



A microscopic cedar pollen grain found on Baby Doe's clothing was one of the indicators that pinpointed that the unidentified little girl was from the Boston area. Image courtesy of Andrew Laurence, Customs and Border Protection.

In an episode that seemed straight out of a television show, forensic scientists were able to determine that "Baby Doe" had lived in the Boston area by examining traces of pollen on her clothes (Laurence and Bryant, 2019), which allowed the police to focus on that city and eventually identify her. It turns out that every area has its own unique "pollen fingerprint," allowing scientists who study pollen—palynologists—to deduce where clothing, drugs, or even explosive devices have originated. This is only possible because of the existence of pollen collections from many different areas across the country and around the world, against which pollen samples can be compared.

and informal learning (see Chapter 3). Without biological collections, educators would lose an exceptional resource for training generations of scientists as well as enhancing both scientific and STEM literacy (Cook et al., 2014; Lacey et al., 2017; NASEM, 2016, 2018d). Integrating the use of biological collections into formal and informal education builds competencies in applied and pure research, data collection and analysis (data literacy), and core biological principles. Moreover, collections introduce students and early-career scientists to extensive and readily available resources that they can explore and use to innovate and develop new lines of inquiry.

One way to understand the value and promise of biological collections is to envision what could happen if there were not a renewed and expanded commitment to maintain the diversity of biological collections and promote their use. Significant domains of basic and applied research would certainly be hindered. Living collections, because of the nature of their maintenance, are particularly vulnerable to inconsistent preservation and, as such, would be irreparably damaged. The loss of genetic stock collections, each a centralized source of materials for a global research community, would irreparably sever the connection between past, present, and future research needs of thousands of research labs that rely on them. Researchers would have to revert to peer-to-peer exchanges, which would greatly hamper long-term availability and quality control, and many advances that cannot even be imagined today would never be made.

Connecting Biological Collections to Create Broad Impacts

There is a growing recognition that integrated global initiatives that apply diverse perspectives, institutions, and resources to prevent and respond to issues of high international priority such as emerging infectious disease, biodiversity loss, food security, invasive species, or climate change are a key approach to achieving an effective and lasting response (Cunningham et al., 2017; Johnson et al., 2011; Machalaba et al., 2015; Myers, 2018). If they were more fully connected across diverse disciplines, biological collections could play a much larger role in these initiatives (Dunnum et al., 2017). Biological collections can provide a platform with which to examine facts, deepen knowledge, and generate innovative solutions to these emerging challenges. More than ever, the community of users could take advantage of the biological collections infrastructure to develop a flexible, distributed, and coordinated network of biological and informatics resources to address research and educational mandates. For instance, biological collections could provide valuable, irreplaceable resources that could contribute to at least six of NSF's 10 Big Ideas (see Box 1-8).[13]

CHALLENGES

Despite their important role as critical infrastructure for research and education and the promise detailed above, biological collections are in jeopardy.

[13] See https://www.nsf.gov/news/special_reports/big_ideas.

They are consistently undervalued and often underfunded. Each year brings new reports of large and small collections threatened with budget cuts or closure (Deng, 2015; Lambert, 2019b). The frequency of such reports provides evidence of a growing issue that needs the immediate attention of scientific decision makers and funders alike. In spite of the broad and varied nature of biological collections, the committee identified many common issues, opportunities, and challenges faced by all. Several of these challenges are related to funding in one way or another. But if one looks beyond this basic issue and asks why funding is such a problem, other challenges emerge. Many of these fall under two broad categories: a lack of recognition of the value of collections, and issues with coordination, integration, and accessibility.

Challenge: Lack of Recognition of the Value of Collections

A consistent challenge facing biological collections is a lack of awareness of the value of these collections to scientific research, innovation, and education and missed opportunities to take advantage of this key infrastructure. Despite the rich history of research, discovery, learning, and innovation built on biological collections, they remain a treasure trove of untapped knowledge because both their contribution and importance are often not widely appreciated or fully comprehended. Natural history collections have been falsely regarded as drawers full of quaint but irrelevant old specimens by some, but well-curated collections contain a temporal record of specimens that have been studied and annotated by generations of scientists. Such collections need to be actively growing, embracing new kinds of specimens, and adopting new technologies to extend their value. There may also be a misconception that the use of "classical" or living model organisms is waning (Hunter, 2008; Jarrett and McCluskey, 2019). In fact, in the past decade, there has been a surge in the distribution of model organisms by living stock collections, which are now offering new materials such as genomic DNA, arrayed strains, and insertion or disruption mutant strains or libraries generated using targeted mutation techniques such as CRISPR. The value of biological collections could be made clearer through targeted initiatives with experts in education, policy, and communication. Ultimately, the collections community needs to improve its ability to communicate the importance of specimens in research and education to a wider audience, especially to funders and decision makers.

Challenge: Biological Collections Infrastructure Taken for Granted

Like all scientific advances produced by the research enterprise, the nation's biological collections require robust resources and infrastructure to maintain them. The physical, digital, and intellectual capital of this infrastructure underlies every aspect of management of, and access to, collections. However, the overall infrastructure that supports biological collections and makes them accessible to the research and education communities is, at best, underappreciated and, at worst, ignored—often at their collective peril. Many funders simply fail

BOX 1-8
Examples of How Collections Contribute to the National Science Foundation's Big Ideas

Six of the National Science Foundation's (NSF's) 10 Big Ideas (with brief descriptions from the NSF website in italic) are linked to the chapters in this report that describe how collections contribute to the Big Ideas.

- **Growing Convergence Research:** *The grand challenges of today—protecting human health; understanding the food, energy, water nexus; exploring the universe at all scales—will not be solved by one discipline alone. They require convergence: the merging of ideas, approaches, and technologies from widely diverse fields of knowledge to stimulate innovation and discovery.* Chapter 2 of this report presents a range of opportunities that garner the power of convergence through transdisciplinary research using specimens and their extended data.
- **Understanding the Rules of Life: Predicting Phenotypes:** *Elucidating the sets of rules that predict an organism's observable characteristics, its phenotype. Life on our planet is arranged in levels of organization ranging from the molecular scale through to the biosphere. There exists a remarkable amount of complexity in the interactions within and between these levels of organization and across scales of time and space.* Chapter 2 of this report provides the past, present, and future contributions of living and non-living collections to fulfill this goal.
- **Mid-Scale Research Infrastructure:** *Developing an agile process for funding experimental research capabilities in the mid-scale range. The National Science Foundation's science and engineering activities rely increasingly on infrastructure that is diverse in space, cost, and implementation time—everything from major observatories to nationwide sensor networks to smaller experiments. There are many important potential experiments and facilities that fall between these; this gap results in missed opportunities that leave essential science undone. The long-term consequences of that neglect will be profound for science as well as for our nation's economy, security, and competitiveness. We need a new approach to research infrastructure, one more dynamic and flexible in*

to recognize the importance of making a long-term commitment to the infrastructure that is needed to maintain, grow, and make biological collections available, in much the same way that oceanographic research vessels support ocean science. Combined with a scarcity of funding, the lack of a long-term commitment or plan for this infrastructure (see Chapter 4) creates a situation where funding for biological collections is often insufficient and unpredictable. As discussed in detail in Chapter 4, priceless and irreplaceable research materials and records of the world's biodiversity are at great risk from everything from outright disaster and federally mandated shutdown to the simple failure of environmental control systems. Changing institutional priorities can be equally *response to this new reality*. This report as a whole describes how biological collections are an essential element of the life science research infrastructure (see Chapter 4).

- **Harnessing the Data Revolution:** *Engaging NSF's research community in the pursuit of fundamental research in data science and engineering, the development of a cohesive, federated, national-scale approach to research data infrastructure, and the development of a 21st-century data-capable workforce*. Chapter 5 of this report describes the important ways digital data are used to benefit research in yet unimaginable ways.
- **Navigating the New Arctic:** *Establishing an observing network of mobile and fixed platforms and tools across the Arctic to document and understand the Arctic's rapid biological, physical, chemical, and social changes. Current Arctic observations are sparse and inadequate for enabling discovery or simulation of the processes underlying Arctic system change or to assess their environmental and economic impacts on the broader Earth system*. Chapters 2 (innovative and transformative specimen-based research) and 8 (community collaboration) lay the foundation for understanding the critical role that collections play in understanding and documenting changing conditions in the Arctic (e.g., Colella et al., 2020; Hoberg et al., 2013).
- **NSF Includes:** *Transforming education and career pathways to help broaden participation in science and engineering. The program's structure will provide a networked testbed for research on STEM [science, technology, engineering, and mathematics] inclusion. This will enable participants to determine the key components and approaches that lead to progress in STEM inclusion as well as the elements that allow successful local alliances to be scaled up for broader use*. Chapter 6 of this report focuses on workforce and includes diversity and inclusion. A critical component of this effort is the value of biological collections research to a range of demographic and psychographic groups, including tribal peoples, as well as citizen/community scientists contributing to the body of knowledge. Chapter 6 also recognizes within its diversity mandate that STEM education is supported by an ecosystem that includes not only schools and universities, but also museums, community organizations, and afterschool/summer activities.

devastating, sometimes resulting in collections being slowly shuttered or even discarded (see Box 1-9). Every collection that is lost means losing years of work and invested resources as well as a skilled workforce, which could in turn lead to major missed opportunities and a decrease in scientific competitiveness for U.S. researchers (Boundy-Mills et al., 2016). Perhaps the worst loss of all, however, is the lost connection to Earth's rich history of life and the knowledge necessary to address pressing societal challenges. If biological collections are to maintain—and increase—their value to science and society in the coming years, careful attention will need to be paid to enhancing collections for future

BOX 1-9
Biological Collections Around the World in Peril

In the overnight hours of September 2, 2018, a fire rapidly escalated into an inferno in Brazil's Museu Nacional in Rio de Janeiro. In just a few hours, millions of irreplaceable specimens and the research careers of dozens of scientists were destroyed. Writing in an op-ed in the *Los Angeles Times* days after the Rio fire, John McCormack (2018), a professor of biology at Occidental College, cited decrepit infrastructure, poor record-keeping, and skeleton staffs produced by years of budget cuts as among the growing concerns facing museums in the United States. The fire at Museu Nacional is just the latest in a string of high-visibility disasters. From 2010 to 2020, Brazil's biological collections were particularly prone to fire damage (Rodríguez Mega, 2020), but there have been problems in multiple other countries as well, such as the 2016 fire at the National Museum of Natural History in New Delhi (Nijar, 2016). Some of the world's most important biological collections have been struck, underlining the precariousness of their infrastructure.

Besides such physical destruction, loss of funding or personnel have been equally devastating for biological collections. After a series of reorganizations in the past few decades, the New Zealand National Museum Te Papa Tongarewa lost almost half its collection managers and curators, jeopardizing the fate of this collection (McDonald, 2018).[b] In 2019, the governor of Alaska proposed to completely cut the state appropriation to the University of Alaska's Museum of the North in addition to imposing severe cuts on the university's annual investments in research (Lambert, 2019a). In 2020, the coronavirus disease 2019 (COVID-19) pandemic forced nearly all museums in the United States to close their doors (Pennisi, 2020). For example, the American Museum of Natural History in New York City had to cut its staff by 20 percent, furlough an additional 250 staff members, and restrict access to the museum to the remaining staff. The long-term impact of this pandemic was unknown at the time this report was published, but will undoubtedly have consequences. A survey[a] released by the American Alliance of Museums in July 2020 indicated that possibly one-third of museums will not reopen.

---

[a] See https://www.aam-us.org/2020/07/22/a-snapshot-of-us-museums-response-to-the-covid-19-pandemic.

research needs and preparing for the loss of infrastructure or expert workforces through retirement or staff attrition.

Challenge: Clear Metrics to Evaluate Biological Collections

Interest and demand for the clear and robust evaluation of research institutions are rising nationally and globally. However, measuring the impact of the nation's biological collections on research and education is difficult because it requires the same stringent standards expected to produce credible, robust scientific research in general. Biological collections lack the resources—financial support, time, and expertise—to develop and implement evaluation plans and to collect and monitor data and information. In addition, there is no consensus on community-wide standards for evaluation and metrics.

Challenge: Coordination, Integration, and Accessibility

Another category of challenges relates to various coordination, integration, and accessibility issues. Historically, biological collections were developed independently of one another, and they have traditionally operated as independent collections, with relatively little coordination or integration among them. This fragmented nature limits the usefulness of the national system of biological collections, leaving potential users of the system often uncertain about what is available and where they can find materials of interest. A lack of coordination and integration both within and across different collections also hinders research involving multiple collections.

Challenge: Incomplete Inventory of Existing Living Stock and Natural History Collections

The precise number and extent of biological collections in the United States are unknown, in part because there is no system-wide process for identifying and cataloging these collections. The number of biological collections is in flux as new collections are created and existing ones are transferred, combined, and discarded. In addition, there is no mechanism to track either the large number of project-based collections that are housed in individual research labs or privately owned collections (which are not covered in this report), which may be eventually accessioned into larger repositories. The extent or value of those collections is not known. A related challenge is that the data associated with those collections, including images and genetic sequence data (see Chapter 5), will require new bioinformatic resources to digitize (if necessary) and publish the acquired data onto online repositories that are available to the research community.

Recent estimates suggest that there are about 1,800 natural history collections in the United States, representing about one-third of all global collections (Kemp, 2015). The most comprehensive list of natural history collections in the

United States, the iDigBio Collections Catalog that lists ~1,600 collections,[14] is an advance over previous efforts, but it is static and not yet complete. Certain living stock collections have self-organized into federations, networks, and consortia, such as the World Federation for Culture Collections, the United States Culture Collection Network[15] (USCCN), Crop Germplasm Committees,[16] and the International Society for Biological and Environmental Repositories,[17] with a growing number of registered collections. When researching the number of living stock collections for which information is available online, experts on the committee estimated that there is a minimum of 2,855 living stock collections in the United States. However, the number of living stock collections is likely grossly underestimated (e.g., McCluskey, 2017), in part because of the diversity of these collections and the different research communities they serve. For example, there is no central registry of genetic stock collections or biological resource centers,[18] which harbor untapped resources for basic research as well as medical, agricultural, and biotechnological applications (Wang and Lilburn, 2009). To start closing the gaps, the taxonomy group at the National Center for Biotechnology Information has created a platform to connect genetic sequence records to specimens of living organisms preserved in living stock collections and to vouchers—representative specimens stored for later examination—held in natural history collections (Sharma et al., 2018). However, without a comprehensive, systematic, and continuously updated inventory of all biological collections, the ability to effectively address the needs of these collections as a community is severely hindered.

Challenge: Limited Community-Wide Coordinating Mechanisms

Many biological collections in the United States and around the world remain largely disconnected. Often, because of geographic or institutional divisions and a lack of funding or awareness about the value of their research materials, project-based collections are in temporary or even permanent storage, usually in the care of the principal investigator funded for the original research. Under such conditions, these resources are not available to inform the wider research community. On a larger scale, creating a coordinating network, developing a common vision, and communicating the value of a network of biological collections to the scientific community, funders, and society as a whole are hampered by the fact that researchers, curators, collection managers, and users are spread across many institutions and often balance multiple responsibilities. This lack of a common vision directly affects their ability to develop a strategy for preserving, growing, cataloging, digitizing, and using collections. Recent support

---

[14] See https://www.idigbio.org/portal/collections.

[15] See http://www.usccn.org/Pages/default.aspx.

[16] See https://www.ars-grin.gov/CGC.

[17] See https://www.isber.org.

[18] Institutions that store and maintain the subject materials of biological research and provide services related to these materials. They also collect and store data and information relevant to their holdings (Wang and Lilburn, 2009).

by NSF's ADBC program has helped to unite the U.S. natural history collections community, across taxa and geography, in unprecedented ways; however, more can be accomplished.

VISION FOR THE NEXT DECADE

Many publications and contributions of individual experts were invaluable in guiding the work of the committee, particularly in regard to the distinct, perhaps unique, needs of different types of biological collections. The committee's conclusions and recommendations represent the deliberations of its members, who recognize both the challenges and power of a diverse national system of biological collections and the reality that budget issues necessitate trade-offs in programmatic priorities. The committee also recognizes the importance of the historical roles of biological collections while envisioning and expanding new functionalities and capabilities to meet 21st-century needs.

The significance of biological collections as research infrastructure continues to grow in ways that were unanticipated 20 or even 10 years ago. With strategic thinking and steady resource investments, biological collections could continue to be at the heart of scientific advances and education for the foreseeable future. Looking ahead, the committee developed a common vision for how best to support, promote, and utilize the biological collections community over the next decade:

> *Provide long-term support for collections-based scientific research, instill a culture of stewardship for and access to biological specimens, build and grow biological collections to better represent global biodiversity in space and time, promote access to biological collections as important educational resources for the general public, and encourage the exchange of biological resources and knowledge.*

With this vision, the major aim of this report is to stimulate a national discussion regarding the goals and strategies needed to ensure that U.S. biological collections not only thrive, but continue to grow throughout the 21st century and beyond. This expansive endeavor requires creative leadership that encompasses a wide range of perspectives and expertise to identify the needs of collections infrastructure and ensure the collections' sustainability and growth.

How can this vision be realized? In this report, the committee first explores the ways that biological collections have contributed to society by advancing scientific discovery and innovation, enriching education, connecting nonprofessional communities to nature and science, and preserving Earth's natural science heritage (see Chapters 2 and 3). Then the committee addresses how the biological collections community is working toward a common vision in light of today's challenges, recognizing that the future success of the biological collections community—curators, collection managers, directors, and users of biological collections—depends on addressing four interrelated issues:

1. upgrading and maintaining the physical infrastructure and the growth of collections (see Chapter 4);
2. developing and maintaining the tools and processes needed to transform digital data into an easily accessible, integrated platform as cyberinfrastructure increases in complexity (see Chapter 5);
3. recruiting, training, and supporting the workforce of the future (see Chapter 6); and
4. ensuring long-term financial sustainability (see Chapter 7).

Realizing this vision will require enhanced communication and collaboration within the biological collections community and beyond (see Chapter 8). The committee recognizes the lack of a common place where issues that span the collections community can be addressed. For curators, there is no single association or professional society dedicated to creating opportunities for networking, collaborating, recognizing, supporting, and promoting the collective research enterprise that is supported by biological collections. Until recently, convening opportunities have been limited to either particular research disciplines that the collections serve (often taxonomically bounded) or to particular regional settings, which is not conducive to the dissemination of information and resources pertinent to the advancement of specimen-based research and curatorial best practices.

In contrast, the biological collections community has various networks to address concerns about the management, care, and distribution of biological collections. These networks can ease the way to establishing strong guidelines, providing training, developing best practices, and facilitating the use of collections in collaborative research as well as in formal and informal education. Networks also provide a platform for strategic thinking and developing solutions to problems of broad societal importance. For instance, the Society for the Preservation of Natural History Collections has made tremendous progress over the past three decades in building a community-driven organization with a common voice. Certain living stock collections have also been successful in establishing national and global networks, such as USCCN (McCluskey et al., 2016; Wu et al., 2017) and could serve as a model for other biological collections. Collections for which the data are digitized and published as part of such national and international networks can also benefit from services that allow these collections not only to gauge the accuracy and completeness of their data but also to comply with relevant legal requirements such as the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity[19] (Nagoya Protocol) (see Box 1-10).

These are compelling arguments for the creation of a common place to develop a unified vision, exchange ideas, pool resources, and in other ways cultivate a thriving biological collections community. To facilitate the realization of this vision, this report explores and offers recommendations for

[19] See https://www.cbd.int/abs.

BOX 1-10
Navigating International Requirements for Sharing and Exchanging Biological Materials and Data

Adding to the complexity of bridging international endeavors, new domestic and international regulations have set out a strong legal framework on access to and use of genetic resources preserved in both natural history and living stock collections. The Convention on Biological Diversity (CBD) recognizes the importance of preserving global biodiversity and sharing benefits arising from the use of genetic resources. The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity (the Nagoya Protocol), a supplementary international agreement under the CBD, is an emerging challenge affecting researchers and biological collections globally (McCluskey et al., 2017, 2018).

Although the United States is not a signatory to the CBD, U.S. researchers need to follow the Nagoya Protocol regulations if the biological specimens they use originated in signatory countries. Each provider country is establishing its own laws and regulations that detail its rules for accessing specimens and their genetic resources and its requirements for sharing benefits arising from their use.

Country-specific legislation guided by the Nagoya Protocol may require that collections users keep all usage records for acquired collections, including derived publications, patents, and products, and to report these uses to the countries of origin. In addition, collection managers need to confirm that deposited specimens were collected with proper permits and make associated documents available to users. While essential to promoting transparency, this places enormous additional responsibilities on collection managers with little or no extra funding to support the increased cost of implementation.

Benefit sharing can be monetary or in-kind, such as training, capacity building, and collaborative research activities. Penalties for noncompliance, such as fines and the confiscation of research equipment, vary by country (Rochmyaningsih, 2019). Digital sequence information is also under consideration for inclusion in the Nagoya Protocol requirements, although this is currently an unsettled issue. Researchers are having difficulties complying with the Nagoya Protocol requirements (Watanabe, 2017), which is a challenge to access and use of materials and intensifies the need to strengthen both U.S. and international collections.

community-wide, collaborative mechanisms, such as the creation of an Action Center for Biological Collections and the development of a Decadal Plan to guide major investments in the nation's biological collections (see Chapter 8). While collaboration is essential in research, evidence suggests that collaboration dynamics and outcomes vary greatly across institutions, fields, and missions and even in the motives among members of individual research teams in ways that could create barriers to innovation (Bozeman et al., 2013; Katz and Martin, 1997). Along with the biological collections community, professional societies and funding agencies will play a critical role in providing leadership to achieve this vision, which will also require sensitivity to inclusivity to engage the community in ways that ensure all voices are heard.

2

Advancing Discovery, Inspiring Innovation, and Informing Societal Challenges

Biological collections are a critical part of the nation's science and innovation infrastructure. Preserved, fossil, and living specimens constitute a vast repository of biological and ecological data about Earth's biodiversity (Bates, 2007; Meineke et al., 2018a; Wildt, 2000). They provide the foundation for scientific knowledge about past and present organisms, how they are interconnected, and the ways in which their physical and genetic characteristics change over time and space. Specimens and their associated data—from genetic and molecular signatures to digital label data and images—also serve as source material for discovery and hypothesis-driven research across life science. Numerous publications have documented how biological collections underpin basic discovery science such as taxonomy, genomics, systematics, evolutionary biology, and biogeography within and among taxa-focused disciplines (e.g., microbiology, botany, mammalogy, herpetology, ichthyology, and mycology); they also support much of the applied research that drives innovation and provides crucial knowledge about such pressing societal challenges as sustainable food production, biodiversity, ecosystem conservation, and improving human health and security. As new technologies and methodologies in research provide new insights about these specimens, sometimes making possible scientific uses never thought imaginable, the value of biological collections increases even more.

This chapter outlines the fundamental ways in which biological collections support scientific research by preserving biological and ecological knowledge over time and space, enabling new biological discoveries, deepening and widening the scientific understanding of complex societal challenges, and driving scientific innovations. The chapter also touches on best practices for evaluating and consistently measuring the impact of biological collections and how their contributions to science and society continually expand.

A VAST DATA-RICH REPOSITORY

The vast number and types of biological specimens housed in U.S. biological collections make it possible for them to contribute to scientific research in a myriad of ways. For example, biological collections play an important role in providing materials—sometimes unique and rare—that can be studied in various ways, such as by comparing their genomes with information on their phenotypes, distribution, and ecology that can be found in the physical specimens themselves and their metadata. Scientists estimate that 800 million to 1 billion specimens are housed in U.S. natural history collections alone (Kemp, 2015).[1] These combined with living stock collections, which continually propagate and multiply organisms for research, result in a total number of U.S. biological specimens that undoubtedly exceeds 1 billion.

The immense data held in these collections capture a large amount of knowledge about species morphology, biology, traits, and distribution. The use of biological collections and their associated data in research has increased in the past decade in part due to the amount of digital data available online in searchable databases (Ball-Damerow et al., 2019; Hedrick et al., 2020; Nelson and Ellis, 2018) (see Figure 5-1). As described in Chapter 5, this transformation and the increase in the accessibility of digitized specimen data have been so profound that undigitized collections are now referred to as "dark data" by the biological community. Advances in research technologies and methodologies have also been instrumental in increasing the use of biological collections data in scientific research as well as in generating new and valuable types of biological collection data. For example, techniques from genetics, chemistry, physics, and engineering have made it possible for biological specimens to become resources for entirely new fields of research, such as isotope ecology and paleoecology. Among the most prominent sets of new technologies that have expanded data and use of biological collections in research arose from the -omics[2] revolution (see section on Enabling Biological Discoveries below).

As indicated in Chapter 1, the billions of specimens held in biological collections are increasingly accompanied by a rich complement of additional biological material and data (see Figure 1-2) that are being used both to generate new insights about life on Earth and to open new avenues of inquiry in almost every field of science, medicine, and technology (Boundy-Mills et al., 2016; Riojas et al., 2019; Schindel and Cook, 2018; Webster, 2017). A single specimen or series of specimens, if studied by multiple investigators, immediately becomes a nexus that ties disparate studies together. Historically, many biological collections have not included specimens with diverse preparations or broad taxon representation per field sampling event. But biological collections provide a natural platform for data integration, particularly when holistic or "extended specimens" are available that facilitate diverse sets of questions (Hedrick et al., 2020; Lendemer et al., 2020; Schindel and Cook, 2018; Thiers et al., 2019; Webster, 2017). For example, collections of insects and ear punches of rodents now being assembled by the National Ecological Observatory Network would have much greater utility and impact if they included whole specimens with a full complement of associated symbionts and additional taxonomic groups to enable a greater variety of research questions (Cook et al., 2016). This is analogous to

[1] The Integrated Digitized Biocollections is the most comprehensive listing of natural history collections in the United States and lists 1,600 natural history collections in the United States associated with 729 different institutions. This list is incomplete and particularly underrepresents small, regional collections and private collections. See https://www.idigbio.org.

[2] A rapidly evolving, multidisciplinary, and emerging field that encompasses genomics, epigenomics, transcriptomics, proteomics, and metabolomics.

a genetic stock center that integrates strains, disruption or insertion into mutant libraries, genome sequences, and genome annotations using model organisms. Over time, a diverse set of disciplines, technologies, and questions can be joined through individual specimens or sets of specimens, which in turn provides primary biodiversity infrastructure for multiple disciplines. A single specimen is thus transformed into an extended specimen that includes the physical specimen itself and any derivative products. This makes it possible for interdisciplinary researchers to study interactions among organisms, communities, and species (Schindel and Cook, 2018) and leads to a new understanding and appreciation of the vast data-rich biological collections repository.

FUNDAMENTAL WAYS IN WHICH BIOLOGICAL COLLECTIONS SUPPORT SCIENTIFIC RESEARCH

Biological collections facilitate research on diverse taxonomic, temporal, and spatial scales. Traditionally they have been most heavily utilized by researchers trying to classify and understand the origins of biodiversity, including terrestrial and marine species as well as microbes. Increasingly—due to a myriad of factors including increased digital access to collections and changing technologies in the biological, physical, and chemical sciences—collections are being used by researchers across the scientific spectrum, to answer diverse questions of immediate relevance to society. Collections provide the raw data for tracking pathogens, identifying invasive species, and many other pursuits that require real-time monitoring. The following sections highlight some of these diverse research agendas, with the aim of outlining the centrality of collections for scientific inquiry and verification with physical specimens.

Preserving and Expanding Knowledge

Each specimen is a unique, tangible, and often irreplaceable representation of life on Earth—past and present. Biological collections maintain specimens of every species known, both "type specimens," which are the specimens originally used to describe a species, and other specimens subsequently collected over time during various explorations and recording events. Sometimes a single natural history or living specimen is all that is known about a species, but these specimens contain the genetic benchmarks and baseline data against which all modern observations and experimentations can be compared. More generally, biological collections serve as the primary source of research material for studying species as well as the main source of information about species, including information about their genetic material, geographic ranges, and morphological characteristics—all of which is used to define the basic units of life on Earth along with their evolutionary histories, their distributions, and the processes that gave rise to them. For example, biological collections are indispensable for exploring and investigating biodiversity and species conservation and for providing a temporal window—on the order of decades, millennia, or even geological epochs—into environmental change

(Meineke et al., 2018a). The primary focus of natural history collections, and to some extent biodiversity living collections, has for centuries been taxonomy,[3] species delimitation, and comparative biology (NAS, 2005), while the main goal of living stock collections has been to allow researchers from varied disciplines to build on knowledge about basic biological functions (McCluskey et al., 2017), although the distinction has not been absolute, and biodiversity living collections have also been used for taxonomy, species delimitation, and comparative biology.

While the spectrum of possibilities has been greatly increased thanks to technological advances in various areas, from curation to digitization (see Chapter 5), the core utility and the organization of natural history collections remain heavily influenced by the original emphasis on biodiversity discovery. Many researchers and museums focus on gathering collections of species found regionally, whereas others aim for comprehensive global collections that contain all the species of a given group of organisms. For instance, some biological collections, such as the ornithology collections at Tring in Hertfordshire, United Kingdom, or at the American Museum of Natural History in New York City, contain approximately 95 percent of the known fundamental taxonomic units found globally. The global coverage of other groups of organisms with known abundant species richness, such as insects or microorganisms, is generally not nearly so complete in biological collections, however. On the other hand, most natural history collections maintain a regional focus and, as such, document genotypic and phenotypic variation in specific localities. As local ecosystems are modified and sometimes destroyed, biological collections become the only remaining representations of endangered species that may be driven to extinction, making the specimen information these collections contain essential for biodiversity conservation efforts. For example, a recent National Academies report on the taxonomic status of the endangered red wolf and Mexican wolf reviewed many studies using morphological traits as well as genetic analyses of specimens, many of which are housed in natural history collections (NASEM, 2019a). Like natural history collections, biodiversity collections of living organisms, which consist of independent, wild-type isolates maintained as living organisms, tissues, or cells, are critical for "the ex-situ conservation of components of biological diversity"[4] through perpetual organism replication and the cryopreservation of germplasms. Some of these collections trace their origins to research collections of one or a few investigators, while others are created through the effort of specific research communities. While living stock collections often represent just a sliver of the existing biodiversity, they still serve as a taxonomic resource (Boundy-Mills et al., 2016; McCluskey et al., 2017) as well as providing diverse model organisms and the base material for physiological, biochemical, and molecular studies (Jarrett and McCluskey, 2019; Riojas et al., 2019). For example, the fruit fly *Drosophila melanogaster* has been used as a model organism for genetic research by different research disciplines since Thomas Hunt Morgan used it to elucidate the role that chromosomes play in heredity, for which he was awarded the 1933 Nobel

---

[3] The discovery, description, and documentation of species, the foundational unit of biodiversity.
[4] See Convention on Biological Diversity, art. 2, https://www.cbd.int/convention/articles/?a=cbd-02.

Prize (Nobel Prize Media, 2019); since Morgan's time, studies in *Drosophila* of genetics, physiology, and microbial pathogenesis have resulted in eight additional Nobel Prizes (Rubin and Lewis, 2000).

Technological innovation will continue to increase our ability to extract information from samples and expand our knowledge by addressing questions that were not even envisioned when specimens were originally collected, just as specimens collected centuries ago are today used in new ways, such as for genomics studies, which would have been unimaginable at the time of collection. For this to happen, specimens need to be collected with a more diverse set of research objectives in mind, from stable isotopes and transcriptome and epigenetic studies to host–parasite interactions, microbiome diversity, and the dynamics of biological communities. To future-proof this critical infrastructure, the biological collections community needs to engage with diverse research communities to gain an understanding of the best strategies and priorities for sampling contemporary biodiversity to build collections with maximum utility in the future. Given the existence of sampling biases in today's biological collections (Nekola et al., 2019), it is crucial that future sampling efforts address these biases by coordinating across institutions to both get maximum use from their existing specimen resources and design future fieldwork to maximize temporal comparability and future research impact. The use of biological collections to estimate demographic trends is clearly an emerging area of collections-based research, and in the future, a major goal will be to make this estimation more reliable and accurate, including for common species that can serve as indicators of rapidly changing environments.

Enabling Biological Discoveries

Biological collections are vital assets of the nation's science and technology enterprise and form the foundation for scientific discoveries about the living world around us. Taking advantage of scientific and technological advances, biological collections have the opportunity to make fundamental contributions to science and to inspire people to engage with a new age of discovery. Both physical specimens and genetic repositories of DNA, tissues, and other materials are sources for genomic research, which focuses on the structure, function, evolution, and mapping of genomes for many purposes, including medical diagnosis, agriculture, industrial biotechnology, forensic biology, and conservation. When augmented with collections-associated data such as spatial or phenotypic information and coupled with powerful advances in genetics, informatics, automation, and artificial intelligence, -omics analyses using living and natural history collections can increase our understanding and improve our stewardship of Earth's biodiversity.

Biological collections have played a critical role in providing a wide variety of materials for the development and fine-tuning of new -omics technologies such as genomics, proteomics, and metabolomics, which in turn benefit many fields of research. For example, as described in Chapter 1, since the discovery of the enzyme *Taq* polymerase in a bacteria strain deposited in the American

Type Culture Collection (ATCC) in the 1960s (see Box 1-2), the advancement and accessibility of next-generation sequencing technologies have rapidly transformed life science research by providing the ability to rapidly analyze and profile genomes. Advanced -omics technologies include sensitive molecular biology techniques that allow researchers to obtain results from smaller amounts of DNA from specimens. Successful barcoding by Sanger sequencing has been commonplace for more than two decades, especially for old specimens with degraded or fragmented DNA. More recently, next-generation sequencing, especially short-read technology and sequence capture of targeted genes, has expanded the scope of DNA-based phylogenetic and functional studies and is enabling the inclusion of thousands of species in a single analysis, with samples obtained from natural history collections (Kates et al., 2018, 2019). For example, regulatory regions associated with the loss of flight in birds have been revealed through the genome sequencing of natural history specimens coupled with functional genomics and the analysis of phenotypic traits (Sackton et al., 2019). Also, biological collections were the source of the specimens used for the first sequencing of the Neanderthal genome,[5] and decades-old slides from such collections offered crucial clues about human malarial evolution (Gelabert et al., 2016); in both cases, biological collections were of great benefit in improving our understanding of human evolution and adaptation. Biological collections also have an important role to play in providing materials—sometimes unique and rare—that are used to connect genomes to information about phenotype, distribution, and ecology contained in the physical specimen and its metadata.

Living stock collections provide a vast quantity of high-quality living and preserved specimens that can be used to ensure reproducibility and replicability in science through the long-term preservation of genetic identity (NASEM, 2019d). Decades of research on generations of these living collections have led to fundamental discoveries in basic life science, from cellular and molecular biology or biochemistry to neuroscience or physiology and to applied life science such as new biotechnologies, biomonitoring, or medical imaging. For example, aspects of the cell cycle were identified from the study of the bacterium *Escherichia coli* and the yeast *Saccharomyces cerevisiae* maintained in culture collections (Campos et al., 2018). Indeed, living stock collections provide essential research model organisms used by many scientists, some of whom have been awarded Nobel Prizes in recognition of life-changing discoveries in physiology and medicine (see Box 2-1). Living stocks such as *Drosophila* stocks also support a broad range of genetic and evolutionary research, with emerging uses in behavioral neuroscience and circuitry, non-coding RNA biology, biosensors (Bellen et al., 2010; Rubin and Lewis, 2000; Wangler et al., 2015), and functional genomics (Mohr et al., 2014).

The development of gene editing methods such as T-DNA, CRISPR (clustered regularly interspaced short palindromic repeats), and RNAi to generate knockout or disruption mutations has expanded the range of organisms available for discovery-driven research. The number of model organism species has grown in

---

[5] See https://www.genome.gov/27539119/2010-release-complete-neanderthal-genome-sequenced.

the past decade, with more than 100 species now considered model organisms (Jarrett and McCluskey, 2019). Some organisms maintained in these collections are studied by a specific research community. An example is the squid *Doryteuthis pealei*, which has giant axons up to 1 mm in diameter, enabling neurobiology studies. Other organisms, such as type strains, tissue cultures, or research organisms (mice, zebrafish, non-human primates, etc.) are maintained for their general research value (Jarrett and McCluskey, 2019). Microbial living collections also constitute a repository of biodiversity used globally for cutting-edge research (De Vero et al., 2019). More than one-third of the deposits of microbe strains into patent repositories between 2001 and 2016 were from U.S. collections, and 3 U.S. collections are among the 47 International Depositary Authorities under terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure[6] of the World Intellectual Property Organization (Wu et al., 2018). One is the Department of Agriculture–Agricultural Research Service Culture Collection Northern Regional Research Laboratory (NRRL) Database used extensively for basic and applied agricultural research, such as taxonomy, for biocontrol of plant pathogens, and even for industrial biotechnology. In fact, the existence of this collection was one of the reasons that patent repositories were established: the NRRL collection was the source of the *Penicillium notatum* strain, a discovery that produced economically relevant amounts of penicillin and as such is a foundational collection for the modern biotechnology era. A second one is the National Center for Marine Algae and Microbiota, which holds thousands of species of microalgae maintained as cryopreserved or actively growing cultures. This living collection is tapped for both basic and applied research, especially filling the needs of pharmaceutical, aquaculture, environmental and bioremediation, analytical instrument, and biofuels research (Scranton et al., 2015; Taunt et al., 2018). Finally, ATCC is by far the most used and cited culture collection in the world. Since 1976, more than 99,000 U.S. patents have cited ATCC alone. Many yeast species are used in fermentation processes to produce fine and bulk chemicals, food and feed ingredients, and fermented foods and beverages (Abbas, 2006). These and many other biodiversity collections are used in basic and applied research, including several genome sequencing projects, funded by various institutions, including the National Science Foundation (NSF).

Driving Innovation

The potential for the use of biological collections in transformative and innovative research has never been greater. Beyond the traditional fields of research described above, biological collections have been a major source of inspiration for scientists from other disciplines, such as physics, chemistry, and engineering.

[6] All states party to the Treaty are obliged to recognize microorganisms deposited as a part of the patent procedure, irrespective of where the depository authority is located. In practice this means that the requirement to submit microorganisms to each and every national authority in which patent protection is sought no longer exists.

BOX 2-1
Nobel Prizes from 1958 to 2017 Involving the Use of *Escherichia coli* and *Saccharomyces cerevisiae*

In 1922, an *E. coli* bacteria strain with a short replication cycle that was easy to grow, preserve, and modify was discovered. By 1925 a culture of this *E. coli* strain had been deposited in the strain collection of the Department of Bacteriology at Stanford University, where it was called *E. coli* K-12. This original strain and several mutant derivatives, which can now be found in living collections such as the *E. coli* Genetic Stock Center at Yale University, have played key roles in major discoveries that have been awarded Nobel Prizes:

- 1958: Lederberg: genetic recombination and the organization of the genetic material of bacteria
- 1959: Ochoa and Kornberg: DNA replication, how life copies its genetic code
- 1965: Jacob, Lwoff, and Monod: gene regulation, how genes are turned on or off
- 1968: Holley, Khorana, and Nirenberg: genetic code, the language in which our DNA is written
- 1969: Delbrück, Hershey, and Luria: virus replication, how viruses reproduce inside cells
- 1978: Arber, Nathans, and Smith: restriction enzymes, cellular "scissors" that allow scientists to cut DNA
- 1980: Berg, Gilbert, and Sanger: recombinant DNA, the creation of the first genetically engineered DNA

For example, unconventional uses of collections in the field of synthetic biology and biomimetics—which are explored in this section—emphasize the potential transdisciplinary opportunities that biological collections can help fulfill.

Supporting Synthetic Biology

Living collections have been instrumental in the development of tools—and still provide the founding material—for synthetic biology, an interdisciplinary field that spans biology and engineering. The foundational work in this field was carried out in the microbial model species *Escherichia coli* and *Saccharomyces cerevisiae*. These microbial systems remain central to this field and have been used for complex circuit design, metabolic engineering, minimal genome construction, and cell-based therapeutic strategies (Cameron et al., 2014). Starting in the mid-1990s, DNA sequencing and improved computational tools made it possible to sequence complete microbial genomes. *E. coli* became the synthetic biology workhorse because of how easily its genes are manipulated, its largely documented biology, and its well-studied gene regulatory systems that

- 1989: Altman and Cech: RNA as an enzyme, additional roles for RNA discovered
- 1997: Boyer, Walker, and Skou: ATP generation, how cells make ATP, the energy molecule that powers life
- 1999: Blobel: signal sequences on proteins, one way that cells organize themselves
- 2008: Shimomura, Chalfie, and Tsien: green fluorescent protein, a tag that scientists use to track cell components
- 2015: Lindahl, Modrich, and Sancar: mechanistic studies of DNA repair

In the late 1930s, in what is now the Phaff Yeast Culture Collection at the University of California, researchers discovered a rare wild mutant of *S. cerevisiae* that could be maintained in the lab as a haploid, meaning that it carries one copy of each gene rather than two, while the typical form of the yeast was diploid, carrying two copies of each gene. This mutant property led to use of *S. cerevisiae* as one of the first model organisms and to the following Nobel Prizes:

- 2001: Hunt, Nurse, and Hartwell: how the cell cycle is regulated
- 2006: Kornberg: how genes are regulated during transcription
- 2009: Blackburn, Greider, and Szostak: how chromosome ends (telomeres) are protected from degradation
- 2013: Shekman, Rothman, and Sudhof: how transport and secretion of proteins are regulated
- 2017: Ohsumi: mechanisms for autophagy

provide a convenient initial source of circuit "parts." For example, BioBricks are building blocks composed of either natural or engineered DNA sequences such as promoters, coding sequences, and ribosome binding sites that are used to assemble synthetic biological circuits called devices; a set of devices is then combined to form a system that performs high-level tasks (Knight, 2003). The BioBrick standard biological parts[7] are now used worldwide—for example, at the International Genetically Engineered Machines competition.[8] In addition to *E. coli*, many specimens from numerous biological collections have been tapped to develop BioBricks (Kahl and Endy, 2013; Radeck et al., 2013) and other innovations in synthetic biology. For example, living collections of phototrophic algae, which have a low production cost and use only sunlight to fix atmospheric carbon, are promising candidates for the manufacture of bioproducts, such as biofuels, through genetic engineering or synthetic biology (Wang et al., 2012).

[7] A biological part that has been refined in order to conform to one or more defined technical standards.

[8] See https://igem.org/Main_Page.

Microalgal biofactories have the potential to become sustainable platforms that could produce certain plant-derived products (Vavitsas et al., 2018) and drive the establishment of an algal-based bioeconomy at some point in the future.

Inspiring and Informing Novel Designs

Biological collections provide a largely untapped reservoir of successful solutions to nature's challenges and thus inspiration for biomimetics—the extraction of "good ideas" from nature to solve human problems (e.g., Green et al., 2019). Both natural history collections, including fossils, and living collections are potential sources of innovation, with applications in such areas as textiles, advanced materials, aerospace, electronics, and even wound care through the use of biofilms from living stocks. Earth's diverse species have developed, through adaptations, unique solutions to a wide variety of problems—solutions that are often beyond the human imagination—and human innovators have turned to biomimicry for decades, for example, in the application of animal locomotion to adhesion science (e.g., Autumn et al., 2002, 2014; Peattie and Full, 2007) and in the use of fungi in mathematical studies of fluid dynamics (Roper et al., 2015). Today, there is a new emphasis on biomimicry with the goal of accelerating the transfer between nature and technology by applying direct applications from diverse collections (Green et al., 2019). With billions of specimens in natural history collections worldwide, the phenotypic diversity is immense, and the digitization of these collections is increasing their accessibility for biomimetic work (Hedrick et al., 2020). Particularly relevant are two-dimensional, three-dimensional, and computed tomography images of specimens, while other materials from natural history collections, such as field notes with habitat descriptions, provide the backdrop for understanding phenotypes in the context of their environments. Examples include research on optical biomimetics aimed at improving the performance of reflectors, which has involved the analysis of iridescence in collections of beetles, butterflies, and even the fruits of the marble berry plant (e.g., Diah et al., 2014; Ingram and Parker, 2008; McNamara et al., 2014; Zhang et al., 2014), and also efforts to engineer materials for use under extreme environments, which have incorporated collections of deep-water sponges and corals (Ceballos et al., 2017). Analyses of the integumentary scales of insect specimens using synchrotron small-angle X-ray scattering and electron microscopy have found high structural diversity at the nanoscale, revealing novel polymer and lipid structures with potential applications to biosensing (Forster et al., 2010; Saranathan et al., 2012, 2015; Vukusic and Sambles, 2003). Robotics also takes inspiration from many biological structures and processes made accessible by living and natural history specimens. For instance, biological collections provide diverse resources for the study of bite force and tooth microwear, including studies on humans (Tanis et al., 2018). New partnerships between engineers and the collections community are emerging, with calls from the biomimetic community for increased funding for collections to support fieldwork, for the acquisition of new specimens, for digitization, and for the interpretation of phenotypes and adaptations (Green et al., 2019).

Widening Understanding of Complex Societal Issues

From reconstructing and analyzing important historical changes to direct applications in national security or human and animal health, biological collections are a physical, digital, and intellectual resource that can enable innovation in translational research (Green et al., 2019; Riojas et al., 2019; Wu et al., 2017) for the benefit of science and society. This next section describes research and innovations to which biological collections have contributed, that are informing, and can confidently be predicted to inform, complex societal issues in the future.

Understanding and Forecasting Effects of Global Change

Biological collections are essential to fundamental research on Earth's ever-changing environment (Lister, 2011; Moritz et al., 2008) and on changes in the distribution and diversity of species over time, including research focused on forecasting these changes (Meineke, 2018b). Estimates indicate that 75 percent of terrestrial areas and 66 percent of the oceans have been significantly changed, due primarily to agriculture and food consumption, and that some 690 vertebrate species and 571 species of plants have been driven to extinction in the past 500 years, with an estimated 1 million more extinctions expected by the end of the 21st century (Humphreys et al., 2019; IPBES, 2019).[9] An increasing awareness that Earth is changing has led to calls for rigorous assessments of how these changing conditions, including the loss of biodiversity, will affect the many ecosystem services that humans rely on (Humphreys et al., 2019; IPBES, 2019). Natural history specimens have been referred to as "biological filter paper": as organisms interact with their local environments throughout their lives, they accumulate a record of environmental conditions that can be interrogated through both established and emerging technologies, including chemical, physical, and molecular analyses (see Box 1-1). For example, hormones can be extracted from decades-old natural history collections, making it possible to infer the physiological state of the individuals at the time of capture (Schmitt et al., 2018), and marine macroalgae from herbaria can be processed with new techniques to provide a historical account of ocean conditions (Miller et al., 2020a). As described above, every biological collection specimen represents the occurrence of a unique individual and species at a particular time and location; as such, these specimens provide some of the best windows available into environmental quality and changing conditions (Edwards et al., 2005; Schmitt et al., 2018).

The degree to which collections can enable transformative research, an understanding of changes in biodiversity, and the development of efficient conservation plans depends, in part, on the continuity of the collections in time and space, because having continuous records of environmental and biological changes is important in all of these areas (Bakker et al., 2020). Despite there being more

[9] See https://www.un.org/sustainabledevelopment/blog/2019/05/nature-decline-unprecedented-report.

than 1 billion specimens held in the United States (Owens and Johnson, 2019) for both living and natural history collections, biological collections need to continue growing so that records of changing conditions on Earth can be maintained and extended, and the collecting practices of the collections need to be strategically developed and modified in order to reduce sampling and taxonomic biases in the collections (Nekola et al., 2019) and to provide geographically and temporally comprehensive baselines of biodiversity on which future studies can be based (Bakker et al., 2020; Schindel and Cook, 2018). More than simply establishing baselines in the recent past to understand changes in today's world, collections also provide windows into change in the past, including how ecosystems and societies have adapted and evolved, or not, when faced with change.

One way in which biological collections are being used to develop more complete and effective records of change can be seen in the way that regional hubs organize continual surveying and re-surveying across the nation and across the globe. For example, the Grinnell Resurvey Project, conducted by scientists at the University of California, Berkeley, Museum of Vertebrate Zoology, has documented substantial changes in elevation, abundance, body size, and distributional range of diverse vertebrates in Yosemite National Park and other sites in California, based on comparisons of species ranges inferred from specimens collected 100 years ago with specimens from the past decade (Moritz et al., 2008; Riddell et al., 2019; Rowe et al., 2015). Likewise, herbarium records have documented extensive changes in flowering time associated with increasing global temperatures, even on local scales, such as in the Boston, Massachusetts, area during the past century (Primack et al., 2004). Investigations using U.S. and international museum collections and private collections were the first to demonstrate how species respond to climate change by shifting locations, adapting to new conditions, or experiencing local extirpation (Parmesan, 1996).

Natural history collections, whose specimens range from fungi to dinosaurs and from bacteria to sequoias, are like libraries that chronicle the history of life on Earth. The more than 1 billion specimens in U.S. collections span the globe and provide a window into the past through both paleontological collections and collections of living specimens collected over the past three centuries (Owens and Johnson, 2019). These latter collections provide a veritable time capsule for the study of adaptation, response to climate change, and more. Notably, information about the occurrences of fossil marine taxa extracted from specimen-based literature was the basis for the identification of the five mass extinctions in Earth's history (Raup and Sepkoski, 1982). Paleontologists have used collections of fossil specimens to examine how organisms have responded to past climate change (e.g., Peppe et al., 2011; Saupe et al., 2014, 2015). By providing records of historic and contemporary species distributions, records tied to geographic localities can be used for ecological niche modeling. In addition, preserved samples can be examined using new technologies to explore environmental tolerances. Collectively, biological collections can help forecast how individual species will respond to changing conditions in the future (Humphreys et al., 2019; IPBES, 2019; Schmitt et al., 2018; Tollefson et al., 2019).

Monitoring Change in Environmental Quality

Biological collections play a critical role in providing clues for environmental health studies,[10] allowing closure of the gaps between evidence of exposure to contaminants and regulations. Chemists, particularly those interested in public health, pollution, toxins, heavy metals, and recent environmental change, find abundant uses for biological collections (Ławniczak et al., 2020; Schmitt et al., 2018). This is exemplified by the creation around the world of environmental specimen banks, which provide crucial data for contaminant monitoring, prioritization, and environmental research (Becker and Wise, 2006; Odsjö, 2006; Tanabe, 2006). An example that still makes headlines is the concerning presence of mercury deposition in fish. Varying levels of mercury contamination can be evaluated by comparing archived specimens in natural history collections with contemporary specimens, and this can, in turn, be used to inform policymakers (EPA, 2002; Stoner, 2002). Animals such as raptors (birds of prey, owls, and scavengers), canaries, or fish are known to be excellent sentinels of local environmental quality, including the presence of contaminants (Rabinowitz et al., 2009; Vo et al., 2011). Soot deposited on bird specimens, for example, has been used to track the rise and fall of atmospheric black carbon over the past 135 years (DuBay et al., 2017), while changes in the level of organic mercury have been tracked for more than a century by measuring mercury levels in the feathers of historical albatross specimens (Vo et al., 2011). Similarly, half a century ago a retrospective study on eggshell thickness from archived samples of bird eggs indicated a marked decrease in shell thickness coincident with the onset of widespread dichlorodiphenyltrichloroethane use (Hickey and Anderson, 1968; Ratcliffe, 1967) (see Box 5-1), and this finding led to rapid policy changes in the use of chemical pesticides and herbicides. In short, collectively, biological collections are a valuable resource for the biomonitoring of contaminants over time and space.

Ensuring Food Security and Crop Management

Food security is a major global challenge that will become even more acute as the human population exceeds a projected 9 billion by 2050 (UN DESA, 2019), with the estimated demand for food rising by 70–100 percent (Valin et al., 2014). Compounding this increasing need will be changing climatic conditions that will limit food production in regions where crops are currently grown (Lobell et al., 2011; Scheffers et al., 2016; Vermeulen et al., 2018) and that may allow new agricultural pests to become established and persist. Efforts in plant breeding, plant pathology, and pest control have long relied on biological collections—herbarium specimens, seed banks, entomological collections, crop and livestock germplasm collections, and living stocks of bacteria and fungi—for crop improvement and disease control and prevention and will continue to do so in novel ways. A mainstay of crop improvement, whether for increased

[10] The study of factors in our environment that can affect human health and disease.

yield, drought tolerance, disease resistance, or production in new regions, is the incorporation of wild germplasm through breeding programs with closely related wild species (e.g., Ford-Lloyd et al., 2011; Warschefsky et al., 2014). Herbarium records provide information on where these wild relatives occur and are used to develop expeditions for collecting new wild germplasm (Ramírez-Villegas et al., 2010, 2020). In some cases new germplasm, discovered through herbarium collections, can lead to cultivar improvement worth millions of dollars per year, as was the case, for example, with a new tomato hybrid (NatSCA, 2005). As climatic conditions change, cultivars may no longer be suited to regions where they are currently grown, and new assessments matching cultivars with locations will be needed. Ecological niche modeling using a combination of crop locations and crop herbarium specimens will be important for predicting where crops may best be suited in the future (e.g., Aguirre-Liguori et al., 2019; Vincent et al., 2019). Moreover, modeling that incorporates digitized herbarium data for crop wild relatives may aid in the selection of new germplasm for helping crops meet the challenges of a changing climate; wild relatives that offer greater drought tolerance or adaptation to higher temperatures—identified through analyses based on herbarium records—may be especially valuable as breeding sources for new crops.

Managing Crop Pathogens and Pests

Biological collections are also important for identifying, tracking, and managing crop pathogens (Ristaino, 2020; Salgado-Salazar et al., 2018). Emerging plant pathogens, while always a threat to food security, are an increasing concern in today's world, particularly as climate change alters the conditions under which potential pathogens interact with crops. In some cases, the disease agents are not clear, and comparisons with fungi, bacteria, and viruses held in living stock collections are necessary to identify the cause of a disease and to develop treatments and eradication measures. Tracking the spread of plant pathogens has, in some cases, involved the use of plant and fungal herbarium specimens as sources of fungal or bacterial pathogens (Ristaino, 2020). For example, citrus canker, caused by the bacterium *Xanthomonas axonopodis*, is a serious disease of citrus trees. Using herbarium specimens of infected citrus trees, Li et al. (2007) identified extensive genetic diversity in the pathogen, traced the spread of the disease, and cautioned plant quarantine agencies about the persistence of local genotypes. Natural history observations, gained in part through biological collections, have been key to the development of successful integrated pest management and biological control (Tewksbury et al., 2014), which in turn have resulted in increased crop yields (Pretty et al., 2006).

Improving National Safety and Public Health Capabilities

Because estimates indicate that nearly 75 percent of all newly emergent pathogens in humans are from wildlife (Jones et al., 2008), specimens can play a primary role in mitigating zoonotic diseases. Biological collections contribute unique and invaluable insights to the study of pathogens for humans, animals, and plants by providing a vast library of diverse samples for pathologists, disease ecologists, and epidemiologists. Importantly, collections can help researchers fundamentally transform how they approach emergent diseases, from the purely reactive measures that are now normally employed after a pathogen emerges to a more predictive framework that will make it possible to forecast future emergence and associated epidemics (Brooks et al., 2019; Glass et al., 2006; Kutz et al., 2004; Morse et al., 2012). As the frequency of disease outbreak increases (Smith et al., 2014), due in part to human alterations of ecosystems and wildlife trafficking (Johnson et al., 2015; Karesh et al., 2005; Myers et al., 2013), the contribution of archived and newly collected biological collections is becoming critical to national security and global economies. Estimates of the cost of the 2003 severe acute respiratory syndrome (SARS) outbreak alone range from $5 billion to $50 billion (Pike et al., 2014), but the coronavirus disease 2019 (COVID-19) pandemic, produced by SARS coronavirus 2 (SARS-CoV-2), already has taken a much greater financial and human toll in the United States (Schwartz, 2019) and worldwide.

With their associated databases, collections critically tie discoveries of new pathogens to permanent host specimens and, in turn, to a series of bioinformatics resources (e.g., GenBank and geographic information system applications) that allow for more robust exploration, identification, tracking, and public health responses to zoonotic pathogens (Dunnum et al., 2017). At the time of the 2001 anthrax attack in the United States, specimens collected decades before allowed researchers from the Centers for Disease Control and Prevention to quickly identify the strain involved in the attack (Hoffmaster et al., 2002). Collections facilitate identification and knowledge of the distributional limits of the reservoirs, vectors, and pathogens in addition to their surveillance over time. As climate change transforms global environments, disease dynamics and pathogen distributions will change (Kraemer et al., 2015), and a robust biodiversity infrastructure will be needed that is spatially broad and temporally deep in order to interpret emergence under these newly evolving conditions. Collections provide an essential baseline for monitoring and understanding the dynamics of diseases caused by pathogens carried by mosquitoes, ticks, fleas, snails, bats, or rodents and other organisms (Anderson et al., 2001; Durden et al., 1996; Yanagihara et al., 2014; Yates et al., 2002).

Culture collections provide a critical and robust platform with which to preserve newly emergent strains and also distribute materials in response to public emergencies, including providing the tools needed to diagnose and control diseases. For example, the 1918 influenza strain, which was originally thought to be of avian origin, was subsequently found to be similar to contemporary swine influenza strains (Fanning et al., 2002; Taubenberger et al., 1997), which directed researchers to effective countermeasure strategies (Ferguson et al., 2003). In response to the COVID-19 pandemic, the Biodefense and Emerging Infections Research Resources Repository (BEI Resources) added to its catalog the first clinical isolate from a patient in the United States along with its genomic RNA, recombinant proteins, and quantitative synthetic RNA for diagnostic assay development and validation. These reagents complement 90 coronavirus-related items available for distribution worldwide to allow researchers to develop vaccines, treatment options, antivirals, and diagnostic assays. Humanity's painful experience with COVID-19 has starkly revealed the limits of our knowledge of planetary biodiversity and the urgent need to build more robust biodiversity infrastructure and connect it to public health initiatives.

Understanding Complex Microbial Communities

The microbiome is another area in which biological collections are playing a key role. Both repositories of microbial isolates from diverse microbiomes (e.g., bacteria, fungi, and phages from the Human Microbiome Project[11]) and collections based on the concept of the extended specimen are being examined for microbiome symbionts (Lutz et al., 2017). Microbe and plant collections are also being used in studies of plant–microbe interactions such as the work done by the Phytobiomes Alliance,[12] which aims at improving crop health and productivity (Schlaeppi and Bulgarelli, 2015). Such studies produce large amounts of sequencing data, which show the presence of a large variety of microbes. To further complicate these studies, only a very small fraction of these organisms can be grown in the lab or without the presence of other microbes—and many of them have not even been classified (Cross et al., 2019; Wade et al., 2016). In these cases, the nucleic acid sequences become the sole record of the existence of such microbes, making the databases that store these sequences a new type of biological collection (Alverdy and Chang, 2008). Specimens in microbial collections are also used to generate reference databases for microbiome analysis: thousands of DNA sequences generated from a single sample such as a surface swab or fecal sample are compared with those in a reference database such as UNITE.[13] Curated reference databases consist of DNA sequences, which are linked to species names and collection specimens, from which users can glean relevant information such as the potential for pathogenicity against humans, plants, or animals; habitat range; and tolerance of temperatures, salinity, or osmolarity.

Unanticipated Use of Biological Collections

Technological innovation will continue to increase our ability to extract information from samples and expand our knowledge by addressing questions that were not even envisioned when specimens were originally collected (i.e., serendipity), just as specimens collected centuries ago are today used in new ways, such as genomics, unimaginable at the time of collection. New species

[11] See https://hmpdacc.org.
[12] See https://phytobiomesalliance.org.
[13] See https://unite.ut.ee.

of plants, insects, fossils, and even mammals critical for our understanding of the history of life are discovered in natural history collections, often archived decades before their recognition as a new species (Bebber et al., 2010; Burgin et al., 2018; Fontaine et al., 2012). The same is true for microbial collections. In 2019, 128 historical bacterial collections from ATCC and the BEI Resources catalogs, some almost a century old, were identified using novel technologies. A phylogenetic analysis of sequences from these collections generated major taxonomic changes from the identification of new species and subspecies to numerous re-classifications (Riojas et al., 2019, 2020), thus making these collections useful for future study and demonstrating why long-term sustainability of physical infrastructure is so critical. Although some living stock specimens or their related biological resources may not be frequently used, many collection curators can point to several examples of materials that were at one point deemed of little research use, but later became essential. For example, Zika virus was an obscure isolate in living stock collections that for 60 years was rarely requested until it came to worldwide attention during the Zika outbreak in 2015 (see Box 2-2). Other examples of strains that experienced a surge in use decades after deposit include *Thermus aquaticus* ATCC® 25104™, which harbors a thermostable DNA polymerase (PMID: 5781580; Stern, 2004) at the core of modern biotechnology (see Box 1-2), and *Neurospora* strains in the Fungal Genetics Stock Center (FGSC) collection with the historic *os-2* mutation that confers resistance to fungicides (McCluskey and Plamann, 2008). For such unanticipated discoveries from both natural history and living collections to continue, specimens need to be collected with a more diverse set of research objectives in mind, from stable isotopes and transcriptome and epigenetic studies to host–parasite interactions, microbiome diversity, and dynamics of biological communities. To future-proof this critical infrastructure, the biological collections community needs to engage diverse research communities to understand best strategies and priorities for sampling contemporary biodiversity to build collections with maximum utility in the future.

EVALUATING THE IMPACT

The breadth of contributions to the scientific enterprise and education (see Chapter 3) is one of the major arguments for enhancing and ensuring the long-term vitality of the nation's biological collections. However, that breadth also raises the question of how one can measure the impact of biological collections, documenting what are often invisible or unrecognized contributions, based on very tangible specimens and data. That is, are the collections truly making a difference, and, if so, how big a difference?

Many individual biological collections gather various metrics to document their productivity and the extent to which specimens and their associated data are accessed and used by the research community. For example, metrics typically gathered by natural history collections include visits, loans, specimens examined, and orders filled, among others (see Box 2-3). These metrics may be designated as indicators of uses of the collection for research, teaching, or outreach

BOX 2-2
Re-emergence of Viral Diseases: Zika

In 1947, funded by The Rockefeller Foundation, scientists at the Uganda Virus Research Institute in Entebbe, Uganda, were surveying in and around Uganda for unknown or poorly known diseases, among them Rift Valley Fever virus, Mengo encephalomyelitis, and Semliki Forest virus. They also isolated a flavivirus from a monkey in the Ziika Forest in Uganda, described it in a 1952 paper, and deposited it the following year as part of the living collection of microorganisms maintained by the American Type Culture Collection (ATCC). This new "Zika" virus did not raise many eyebrows or much interest at the time; it was not deemed an imminent threat to public health.

Flash forward 60 years to 2015, however, and an outbreak of a mystery illness that caused severe neurological defects in infants exposed in utero suddenly emerged in North and South America, primarily Brazil. After a few false starts in determining the cause of the disease, epidemiologists tagged the Zika virus as the agent responsible for a variety of symptoms including microcephaly, brain malformations, and other birth defects in infants; adults normally experience only a mild infection similar to a low-grade dengue fever.

The Zika epidemic was soon classified as a Public Health Emergency of International Concern by the World Health Organization. As a result, the scientific community's interest in ATCC's strains of Zika virus grew seemingly overnight. It quickly jumped from relative obscurity to one of ATCC's most requested viruses. ATCC (through BEI Resources) continued to culture the virus, refine it for study, authenticate new isolates, and participate in sequencing efforts that allowed scientists to piece together how the virus spread from Africa to Asia and then America (Shrivastava et al., 2018).

The living collection at ATCC allowed scientists to jump-start the process of identifying and characterizing Zika infections, and its researchers quickly lent their expertise and material to scientists around the world. ATCC and BEI Resources support Zika virus research efforts, such as vaccine efficacy testing and the development of detection assays, with an expanding collection of Zika virus reference materials and solutions, including in vivo and tissue culture–adapted strains; genomic and synthetic nucleic acid preparations; host cell lines and reagents; and custom solutions for expansion, titering, and banking Zika virus.

and are often compiled for annual reports to institutional and funding authorities to document short-term activities and for collections advocacy. Some biological collections track and document the use of specimens and their associated biological materials and data through published citations. Specimens in natural history collections and strains in living collections have unique numbers that can be tracked in the literature. In addition, many biological collections require users to acknowledge the collection when publishing, although this mandate is not always followed. For example, the FGSC established an online bibliography[14]

[14] See http://www.fgsc.net/cite.htm.

BOX 2-3
Example Set of Metrics to Document Biological Collection Access and Use

The University of New Mexico's Museum of Southwestern Biology in 2016 Annual Report

	2012	2013	2014	2015	2016	5-Year Average
1. Collection growth (Specimens Cataloged)	25,446	34,772	103,947	129,245	66,334	71,949
2. Loans Out	99	145	241	176	208	174
3. Professional Visitors to the Collections	307	344	248	945	392	447
4. Collection Database Web Site Hits	396,362	**	233,079	585,913	454,998	417,588
5. Outside Pubs Citing MSB Specimens	76	167	147	189	90	134
6. Peer-Reviewed Publications by Staff	77	54	104	80	59	75
7. Graduate Students	42	42	41	27	56	42
8. Graduate Theses/Dissertations Completed	9	7	11	11	11	10
9. Undergraduate Students	76	66	63	57	73	67
10. Grants/Contracts in Force	76	61	61	82	51	66
11. Grants In Force Total Costs	N/A	N/A	N/A	$2,662,014	$6,354,047	$4,508,030
12. Estimated F&A return	$528,950	$410,871	$436,680	379,129	$211,182	$ 393,362

* 1 UNM, 2 outside, NR – not reported

Source: Cook, 2017; See https://digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1010&context=msb_annual_reports

documenting the use of fungal strains, and it directs scientists to cite a published journal article in order to acknowledge the FGSC (McCluskey et al., 2010). There are not yet widely adopted standards and processes for citation, but technology offers some solutions, such as mobile apps and other mechanisms for inputting, viewing, and retrieving information on collections use. Today, living collections and natural history collections have begun to use data aggregators such as Google Scholar to compile research publications that result from collections-based work (Winker and Withrow, 2013). Electronic citation and tracking of digital specimen records, each with a unique identifier, provide attribution to local collections and enable the assessment of short- and long-term impacts both locally and nationally (see also Chapter 5). The Analyzer of Bio-resource Citations[15] of the World Data Center for Microorganisms is a database of publications and patents that cite biological collections and specific specimens (Wu et al., 2017). As of August 1, 2020, more than 145,000 publications had referenced 79,224 microbial strains belonging to 131 culture collections. In addition, more than 42,000 patents had referenced 44,508 microbial strains.[16] The National Center for Biotechnology Information (NCBI) also tracks DNA sequences deposited in GenBank that are associated with specimens from registered biological collections, through the NCBI BioCollections Database. Other citations and attribution

[15] See abc.wfcc.info.
[16] ABC statistics update 2020—8-05 1:53:03 Analyzer of Bio-resource Citations.

systems are in the early stages of development—occCite[17] is one promising example of an online tool that tracks citations of biodiversity collections—but they cannot yet be implemented at large scales. However, the practice and the development of publication requirements from scientific journals on how to broadly implement citations are still in their infancy.

As detailed in the National Research Council report *Furthering America's Research Enterprise* (NRC, 2014c), scientific impact results from multiple processes over time, and identifying the specific metrics necessary to capture that impact requires careful dissection of the goals, timeframes, and outcomes of the research. Measuring the impact of scientific infrastructure, such as the nation's biological collections, may be even more challenging because the collections' purposes, goals, and scale can vary greatly. However, a substantial body of work provides evidence, resources, considerations, and best practices for evaluating and selecting appropriate metrics that could be successfully implemented by biological collections (Guthrie et al., 2013; NRC, 2005, 2010, 2014c).

Evaluation is typically an iterative process that requires advanced commitment and planning. The first step in developing an evaluation plan is to define the *goals* and intended *outcomes* of a biological collection that are fully integrated with the purposes of the evaluation (see Table 2-1). Outcomes may be categorized as short term, midterm, and long term, depending on the estimated time horizons necessary to achieve them. The second step is to develop an evaluation framework. There are a variety of evidence-based evaluation frameworks, each with distinct strengths and limitations (Guthrie et al., 2013; NRC, 2014c). In general, all evaluation frameworks demonstrate the relationships among goals, the available resources (inputs), the planned activities and services, and the intended outcomes. The third step is to develop evaluation questions. These questions relate to various points along the continuum from inputs to the intended outcomes and impacts, and they clarify the scope of the evaluation. Table 2-1 provides examples of evaluation questions that may be important for different components along the continuum from inputs to desired impacts for a biological collection.

Evaluation questions need to produce answers that are measurable. Hence, the fourth and final step of evaluation planning is to identify appropriate metrics—the quantitative or qualitative measurements used in the answers to evaluation questions. Metrics can be measurements of biological collections' processes (e.g., the quantity and amount of external grants, the number of accessions and loans, perceptions of collections efficiency and efficacy) or products (e.g., the number of publications, the contribution to major meta-analyses, the percentage of collections-trained students who chose careers in science, technology, engineering, and mathematics). Assessing the answers to evaluation questions usually requires a mixture of quantitative and qualitative methods, including the analysis of routinely collected metrics data. Some of the most powerful metrics for evaluating biological collections could be qualitative. For example, sentiments about the ease of use of specimen data portals would be important information related to improving access to data for different types of uses. Evaluators often look for

[17] See https://hannahlowens.github.io/occCite.

TABLE 2-1 Key Evaluation Terminology and Example Questions[a]

	Term	Definition	Examples	Evaluation Questions
Process Components	Inputs	The resources needed for program planning and processes.	• Strategic plans • Budget • Specimens • Personnel • Facilities and cyberinfrastructure	• What is the quality of the inputs? • Are the inputs sufficient? • Are the inputs sustainable?
	Activities	The events, services, or functions that take place.	• Strategic planning and evaluation • Collecting and accessions • Distributing specimens • Digitizing and building data portals • Research • Teaching, training, and mentoring	• Are these processes efficient? • Are these processes effective? • Are the activities proceeding as planned? If not, why? • Which activities strengthen collaborative networks?
	Outputs	The direct products of the activities. Outputs can be subdivided into knowledge, infrastructure, or workforce.	• Research-accessible collections • Publications and presentations • Tools, methods, and standards • Databases and data portals • Collections staff professional development	• Which outputs have been produced? • What is the quantity, cost, timeliness, and quality of what has been produced? • Who is the target audience for each type of output?

continued

TABLE 2-1 Continued

	Term	Definition	Examples	Evaluation Questions
Outcome Components	Outcomes	The intended effects on people, communities, or institutions as a result of the outputs. Outcomes can be subdivided by when they are most likely expected to occur: short term, midterm, or long term.	• Meta-analyses (analyses that combine data from multiple studies) • Improved body of collections-based knowledge • Serendipitous discoveries • Expanded network of collections users • Strengthened collections-based research community	• To what extent are the target audiences aware of the outputs? • Have the target audiences used the outputs at least once? Has their knowledge or behavior changed after use? • Are target audiences satisfied with the outputs and accompanying services? • Where has the use of the collection(s) enabled tackling new research questions, making discoveries, finding solutions to challenges in applied research? • Have research networks been strengthened? • Have participants entered the STEM workforce?
	Impacts	The broader changes in communities, systems, or society that stem from the outcomes. Impacts do not directly result from outcomes, but from multiple interacting factors within and outside of a program or institution.	• Improved quality of specimen-based resources for research • Broader participation in STEM • Greater protection of biodiversity and the environment • Prevention and control of human and wildlife infectious diseases • Increased economic competitiveness	• How much have specific observed outcomes contributed to improved research, scientific leadership, human health, environmental protection, or improved businesses? • What evidence demonstrates that the collection(s) contributed to improved quantity and quality of research?

[a] This table indicates that there are two primary pathways to documenting the outcomes and impact of collections: research and education. However, the table focuses primarily on research. Additional discussion of documenting the outcomes and impact of education is provided in Chapter 3.
NOTE: STEM = science, technology, engineering, and mathematics.

sets of metrics, sometimes called *indicators*, to develop more comprehensive answers about the targeted outcomes.

Measuring Comprehensive Impact

Efforts to assess the impact of scientific research are now reaching broadly beyond academia to include *comprehensive impact*, that is, the impact of scientific research on all of human society and the natural environment, including the effects on the economy, health, policy, and society more generally (e.g., Ravenscroft et al., 2017). Although measuring comprehensive impact is difficult and the methods to do so are still in their infancy, the U.S. STAR METRICS program is an example of a platform that may eventually assess the impact of federal research funding on employment, society, and the economy through an analysis of factors such as health outcomes, student mobility, patents, and industry startups (Lane and Bertuzzi, 2011). Other attempts to assess comprehensive impact are also under development.

Biological collections now have an opportunity to learn from new developments in the field of assessment and go beyond usage statistics and measure impact. Given the increasing and diversifying use of collections and the community's newly generated digital assets, this is an excellent time to connect evaluation experts with the collections community to apply evidence-based approaches to assessing the impact and interpreting metrics. Creating spaces and opportunities to exchange ideas and share best practices would facilitate the evaluation process. Moreover, the time is also perfect to develop national goals and desired outcomes and to build a cyberinfrastructure-supported method for the citation and attribution of digital specimen records and for assessing the collective impact of biological collections.

A Community-Wide Vision

Although individual biological collections may vary in their specific goals and desired outcomes, they share the goal of providing effective and impactful access to physical and digital objects for use in research, innovation, and education. Given this shared goal, along with nascent connections among many collections stemming from NSF's Advancing Digitization of Biodiversity Collections program, the collections community has the opportunity to develop a community-wide vision for evaluating its collective impact and how to measure it. The federal Interagency Working Group on Scientific Collections is in the process of documenting outcomes and impact of only federal science collections, based on existing metrics. The federal work could provide important input into a broader effort to evaluate the nation's biological collections. In addition, the collections community can build on the experiences of other networks that have attempted to shape and measure community-wide impact. For example, research on how to achieve change collaboratively has been explored (e.g., Guarneros-Meza et al., 2018; Sullivan and Skelcher, 2002), with possible lessons and benefits for the biological collections community. More specifically, the library special collections and archives community, through professional societies (e.g., the Association of College & Research Libraries of the American Library Association and the Society of American Archivists) have collaboratively developed, aggregated, and leveraged metrics, and their approach can offer guidance to the community-wide process of evaluation for the biological collections community. Tackling metrics as a community would lessen impediments due to limited resources, personnel, and time; allow the community to take advantage of the knowledge of professional evaluators; and shape common outcomes that can be assessed both at individual collections and collectively.

Connect to National Endeavors

The scientific community, in general, is developing approaches to evaluate its performance and impact. As noted above, STAR METRICS is a U.S. government effort to create tools and a data repository to assess the impact of federal investments in research and development. Specifically, STAR METRICS examines the outcomes of federal investments in science on job creation and economic growth. Major efforts are also under way in other countries including Australia, Canada, and the United Kingdom (NRC, 2014c).

Biological collections will need to communicate with other research endeavors that are having the same conversations about metrics. Connecting the conversation around metrics that we hope to spark in this report to larger, broader conversations already beginning to take place across the research landscape has the potential to lead to metrics that can be integrated across biology. Engaging in higher-order conversations about value and impact can help the collections community—and the scientific community at large—use resources more effectively and take greater advantage of public support. Unless the biological collections community participates meaningfully in these larger evaluation schemes, it risks isolating itself by only developing community-specific measures of impact. To the extent that different biological collections develop a set of shared metrics, they will benefit from selecting best practices or exemplars that show biological collections metrics activities that are consonant with the general discussions occurring about the impact of science.

CONCLUSION

Collectively, biological collections allow research to build and expand on decades of scientific advances and knowledge. Biological collections have a substantial legacy in producing a wide range of benefits for research in the United States and the global community. If biological collections are to effectively promote and expand their contributions and impact, it will require ongoing investment, comprehensive planning, and dedicated stewardship. The global collections community, funding agencies (e.g., NSF, the National Institutes of Health, and the Centers for Disease Control and Prevention), and federal natural resource agencies (e.g., the Department of Agriculture) need to create a partnership to implement a coordinated plan to encourage the strategic growth of collections to support all areas of life science research including genetics, cell biology, biotechnology, and synthetic biology as well as a rigorous assessment of dynamic change in planetary diversity, ecosystems, and biomes. Analytical capabilities (both tools and training) to enable transformative research using biological collections and associated data will be needed to ensure that biological collections are rigorously archived to fuel the greatest diversity of new technologies and approaches. Mass digitization and the expansion of innovative digital platforms can broaden the use of collections and engage virtual communities worldwide. To document and monitor such successes, the biological collections community will need to embrace formal evaluations of its impacts through collaborative approaches. Establishing partnerships with professional evaluators and mechanisms to share resources and exchange ideas will be critical for developing the appropriate tools for evaluating the current roles that biological collections play in research and education as well as for strategically expanding those roles in the future.

3

Contributing to Science Education and Lifelong Learning

Biological collections are powerful educational assets for learners of all ages, backgrounds, skills, and perspectives. They provide a tangible platform that can draw people into *lifelong learning*—ongoing efforts to foster, develop, and expand one's knowledge and skills—whether through formal education, employment in science, technology, engineering, and mathematics (STEM), or by pursuing personal interests throughout life. Biological collections are intrinsically multidisciplinary in nature so they can help individuals learn integrative thinking. The use of specimens and their associated data, in educational curricula and informal experiences, can help students and members of the public explore not only biology and biodiversity but also central concepts in science. Such ideas range from the basic principles of the scientific method (e.g., hypothesis testing, verification, replication, and data extrapolation) to methods that help scientists make sense of complexity to the promise and challenges of data-driven discovery.

By facilitating learning across a wide range of disciplines in formal and informal environments, biological collections can deepen subject-matter expertise and stimulate integrative and generative thinking, which can link disciplines from the sciences to humanities and the arts (Balengée and Triscott, 2010; Ho and Cook, 2013; Powers et al., 2014). Educators also leverage biological collections to drive inquiry-based learning[1] in order to improve skills necessary throughout life such as critical thinking, management, data interpretation, and problem-solving (NRC, 1996). Finally, biological collections empower people from all walks of life to connect to and learn about nature (Mujtaba et al., 2018; Soul et al., 2018), building wonder and providing a source of inspiration and appreciation for the natural world.

This chapter outlines some of the historical and contemporary uses of biological collections in STEM education and lifelong learning. It also touches on basic approaches to evaluating and consistently measuring the impact of biological collections on education and learning.

[1] Inquiry-based learning is a student-centered learning and teaching approach in which students' questions (inquiries) and ideas are prioritized—they "pose questions about the natural world and investigate phenomena; in doing so, students acquire knowledge and develop a rich understanding of concepts, principles, models, and theories" (NRC, 1996, p. 214).

INCREASING STUDENT KNOWLEDGE AND UNDERSTANDING IN FORMAL EDUCATION SETTINGS

Biological collections offer a wide range of opportunities to enhance evidence-based approaches in formal STEM teaching and learning. Because biological collections are tangible, they can provide a natural entry point to biology and biodiversity for kindergarten through grade 12 (K–12), undergraduate, and graduate students who may have limited experiences in nature, through the use of high-quality and developmentally appropriate inquiry-based curricula. Students are attracted to these authentic and tangible resources as they engage in the process of scientific discovery and prepare to design and conduct their own research (NASEM, 2015, 2019e; NRC, 2012b).

There are many examples of how biological collections can be incorporated into classroom curricula or as a means to provide research experience: educational kits, classroom visits, field trips, summer camps, online courses, tutorials, blogs, and teacher workshops are a few of the educational tools and programs created by biological collections staff.[2,3] For example, the Arabidopsis Biological Resource Center at The Ohio State University develops and distributes kits to be used in K–12 and undergraduate classroom settings for students to learn about plant biology and topics such as genetics and gene expression, development, inheritance, hormone physiology, biological responses to the environment, and bioinformatics (see Box 3-1). The Culture Collection of Algae at The University of Texas at Austin and the Chlamydomonas Resource Center are examples of living stock collections that offer educational kits. However, developing and distributing living organisms for education tend to be the domains of for-profit biological supply companies, and not an activity of many living stock collections. Many of the universities housing biological collections incorporate the specimens and their associated data into a wide variety of science courses, from introductory classes to advanced directed studies, to enhance lessons about topics such as genetics, physiology, anatomy, adaptation, evolution, biodiversity, and environmental change. Such courses also afford students the opportunity to learn about organisms and organismal interactions (e.g., symbioses, community structure).

Published in 2013, the Next Generation Science Standards (NGSS)[4] are a new set of science and learning standards through which students make sense of data, engage in scientific and engineering practices, and solve problems in context, enabling students to *learn science* by *doing science* (NRC, 2013). Biological collections can be ideal for NGSS teaching, providing authentic, object-based science experiences that actively engage students in science. Integrated Digitized Biocollections (iDigBio),[5] the National Science Foundation's (NSF's) national resource for digitization of biodiversity collections, oversees standards

---

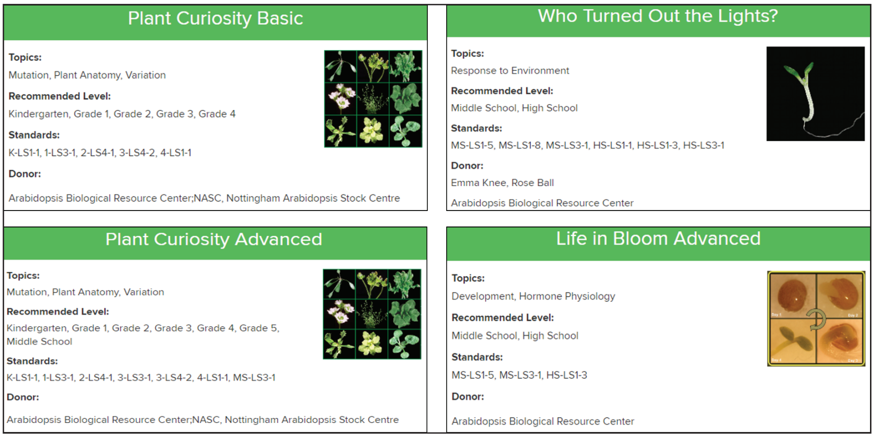
[2] See http://www.usccn.org/methods/Pages/default.aspx.

[3] See http://nscalliance.org/wordpress/wp-content/uploads/2010/01/nsceducate.pdf.

[4] NGSS (https://www.nextgenscience.org) are based on the National Research Council report *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (NRC, 2012a).

[5] See https://www.idigbio.org.

BOX 3-1
***Arabidopsis* in the Classroom**



Top left, top right, and bottom left images courtesy of James Mann, Arabidopsis Biological Resource Center; bottom right image courtesy of Marcelo Pomeranz, Arabidopsis Biological Resource Center

In addition to distributing genetic resources for the research community, the mission of the Arabidopsis Biological Resource Center (ABRC) at The Ohio State University is to "bridge the gap between *Arabidopsis* research and its utilization in kindergarten through college classrooms." With funding from the American Society of Plant Biologists and the National Science Foundation, ABRC's outreach program released 20 education kits designed for use in K–12 and college-level instruction, along with a variety of other educational tools and programs. Six of the kits, known collectively as Translating Research on *Arabidopsis* into a Network of Educational Resources, were developed and tested by ABRC staff. These kits are provided free of charge; most seed stocks are also provided free of charge to K–12 schools. Kits include downloadable materials—specifically, in-depth, ready-to-teach lab protocols and supporting materials, such as instructional videos and datasheets for conducting the outlined experiments. A subset of the available kits has been further developed by ABRC as part of its Greening the Classroom program.

and best practices for digitization and includes an active education and outreach working group. The working group develops and aggregates online resources for K–12 students and educators; many of the educators provide authentic, inquiry-based science experiences that actively engage students in the evidence-based teaching and learning standards of the NGSS. iDigBio also promotes informal science learning through camps for school-age children and develops biodiversity- and digitization-related educational resources for undergraduate students. In this way, efforts to digitize biological collections data through ADBC have catalyzed nationwide opportunities for multiple biological collections to engage students in collections practice and research activities.

Preparing Students for a Data-Driven World

Biological collections are also being used to introduce and develop data science, computer science, and engineering skills (see also Chapter 5). Aligned with one of NSF's 10 Big Ideas, "Harnessing the Data Revolution,"[6] data science is an emerging field important in all subjects and disciplines. A 2018 report by the National Academies of Sciences, Engineering, and Medicine states that all "undergraduates will benefit from a fundamental awareness of and competence in data science" (NASEM, 2018b, p. 1). Biological collections are an exceptional resource for building data literacy at all levels of the data life cycle—finding, generating, curating, evaluating, and using data (NASEM, 2018b). Efforts to digitize biological collections are increasing their accessibility to scientific researchers, educators, and learners. A recent report of the Biodiversity Collections Network notes that "specimen-based data make science accessible through the specimen itself, which is tangible, place-based, and interesting, as well as through aggregated specimen data that are verifiable, relevant, and a logical gateway to data literacy" (Thiers et al., 2019, p. 16).

Two notable endeavors in the biological collections community that have promoted the use of specimen-based data for teaching data literacy are Advancing the Integration of Museums into Undergraduate Programs (AIM-UP!)[7] and the Biodiversity Literacy in Undergraduate Education (BLUE)[8] initiative. AIM-UP! (funded by NSF's Research Coordination Networks in Undergraduate Biology Education, RCN-UBE,[9] from 2010 to 2016) established a network of curators, collection managers, database managers, educators, researchers, and students focused on integrating national history collections into undergraduate biology education. The network spanned 50 institutions in 32 states. Through workshops, professional conferences, webinars, and various social media venues, AIM-UP! built a biological collections data science community that exchanged ideas and generated new approaches to incorporating natural history collections and their

[6] See https://www.nsf.gov/cise/harnessingdata.

[7] See http://aimup.unm.edu.

[8] See https://www.biodiversityliteracy.com.

[9] RCN-UBE is a collaborative program of NSF's Directorate of Biological Sciences and the Directorate for Education and Human Resources. It aligns with an NSF-wide undergraduate STEM education initiative, Improving Undergraduate STEM Education.

associated databases into formal coursework and mentored research experiences (Cook et al., 2014). For example, Lacey et al. (2017) introduced an online, open-access educational module that uses the power of collections-based data to introduce students to multiple conceptual and analytical elements of climate change, as well as evolutionary and ecological biology research. Demonstration education modules, videos, and other examples of ways to incorporate collections into undergraduate education are available online.

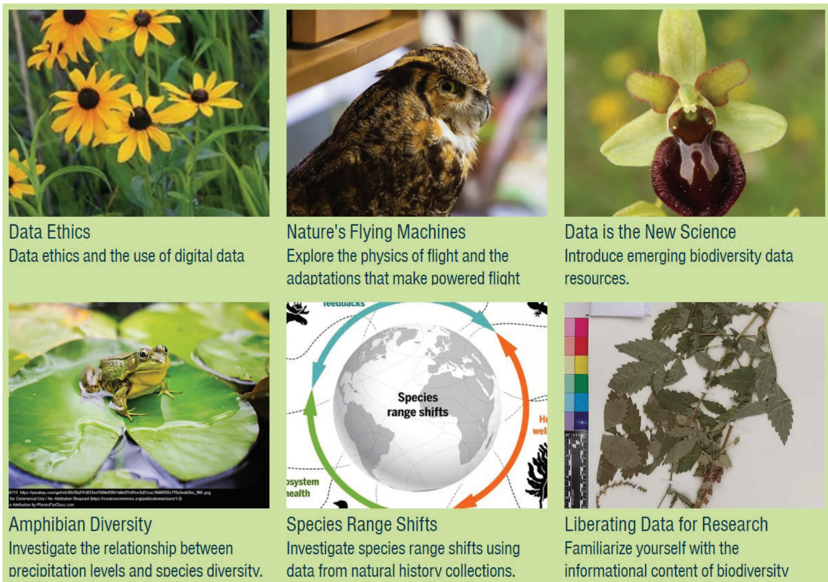
Building from the success of AIM-UP!, NSF funded BLUE to continue to foster a community of biodiversity, data science, and undergraduate education experts and meet increasing need and workforce demands for biodiversity data literacy and integrative analysis skills (Ellwood et al., 2019). BLUE's mission is to define and build consensus around core biodiversity data literacy competencies and also to develop strategies to integrate those data literacy skills and knowledge into introductory undergraduate biology curricula. To that end, BLUE develops exemplar educational materials (see Box 3-2) and actively cultivates a diverse community of practice[10] for undergraduate data-centered biodiversity education through workshops, virtual faculty mentoring networks, webinars, sessions at annual meetings, and invited talks. In its first 2 years, BLUE engaged more than 300 individuals, from undergraduate students to late-career professionals, representing 167 different high schools, community colleges, and universities; 37 different natural history collections; and 22 different collections-associated networks (e.g., iDigBio, the Global Biodiversity Information Facility, and the National Ecological Observatory Network Biorepository).

Enhancing Student Research Experiences

Biological collections can also be used to facilitate synergies between scientific research and education. Education research demonstrates that undergraduate research experiences facilitate active learning and improve biological literacy (AAAS, 2015; Austin, 2018; Bauerle et al., 2011; NASEM, 2017c; NRC, 2015b). The NSF-funded Research Experiences for Undergraduates Program has supported several programs focused on natural history collections including the Academy of Natural Sciences of Drexel University (NSF Award #0353930), University of Iowa Museum of Natural History (NSF Award #15248700), and Field Museum (NSF Award #1156594), among others. Now with digital data from collections, students at universities without a biological collection also have direct access to specimen-based research opportunities (Cook et al., 2014; Monfils et al., 2017; Powers et al., 2014). For example, the RCN-UBE Incubator: Network for the Integration of Natural History Collections in Ecology and Evolutionary

[10] First coined by cognitive anthropologists Jean Lave and Etienne Wenger (Lave and Wenger, 1991), then significantly expanded by Wenger (1998), a community of practice is a group of people who share a concern, a passion about a topic, or a set of problems, and learn how to do their work effectively through regular, ongoing interactions (Wenger, 2000; Wenger et al., 2002). Although the initiation of a community of practice may require funding, effective communities of practice are generative through the value they offer members. As a result, strong communities of practice typically last longer than a project team or task force, continuing as long as they are useful to their members.

BOX 3-2
Select Educational Materials Developed by the Biodiversity Literacy in Undergraduate Education (BLUE) Initiative



Top row left to right: "Black-eyed Susan" by milesizz is licensed under CC BY-NC-ND 2.0; owl image courtesy of Adam M. Sparkes, Central Michigan University Communications; early spider orchid, photo by H. Krisp Bottom row left to right: Bullfrog on lily pad by Jill Wellington; Species Range Shift reprinted from Pecl et al., 2017; Rosaceae, Agrimonia gryposepala by Kathy M. Davis, courtesy of the University of Florida Herbarium, Florida Museum of Natural History

Led by Dr. Anna Monfils of the Central Michigan University Herbarium, BLUE develops exemplar educational modules using data derived from natural history specimens and biodiversity research. For example, "Nature's Flying Machines" enables students to learn about the evolution of flight and the forces that influence flight using digital data from birds and insects. Other modules focus on data science competencies such as best practices to collect, clean, analyze, and present data. As of May 2020, BLUE has published more than 20 open-access modules, 6 of which are shown in the image above.

Biology Course-Based Undergraduate Research Experience focuses on research opportunities afforded by the digitization of collections. The Yeast ORFhan Gene Project[11] is a RCN-UBE program that uses the *Saccharomyces* genome database to integrate researchers (faculty and students) into an effort to assign molecular functions to genes of unknown function in baker's yeast (*S. cerevisiae*), adapting bioinformatic and wet-lab modules for use in classes. Although the needs of research and education are not always the same, student research experiences use the synergies, maximizing investments in collections-based research and education efforts.

Digitized biological collections also make it easier to rapidly respond to an unanticipated disruption to undergraduate biology education. The coronavirus disease 2019 pandemic is driving an unprecedented need for remote learning resources. Of particular concern is the loss of student access to laboratories and field sites that are used for course-based undergraduate research. In response, scientists from Widener University, the Delaware Museum of Natural History, The George Washington University, and collaborators nationwide, with the support of an NSF grant for Rapid Response Research, are developing online course-based undergraduate research experiences using digitized natural history collections.[12]

INSPIRING A LIFELONG APPRECIATION FOR SCIENCE IN INFORMAL EDUCATION SETTINGS

There is abundant evidence across all venues that people learn science in a variety of non-school settings (NRC, 2009). Biological collections are one such important venue and have a history of contributing to lifelong learning and appreciation for science, including sometimes offering opportunities for lifelong learners to participate in science (Prôa and Donini, 2019). This is the case no matter how a biological collection is experienced—through traditional and immersive exhibitions, dioramas, or visual storage methods; through open collection programs for public universities; or during in-depth, out-of-school research internships for middle and high school students (Dawes, 2016; Falk and Dierking, 2013, 2018; George, 2015; Habig et al., 2018; Reiss and Tunnicliffe, 2011; Suarez and Tsutsui, 2004; Tunnicliffe and Scheersoi, 2015). As more specimens become digitized, some natural history collections, such as the Idaho Museum of Natural History, are beginning to offer virtual tours of their biological collections.[13] Virtual tours and online video broadcasts are some of the ways to enable a greater number and diversity of lifelong learners to engage with biological collections.

Biological collections can also inspire awe and stimulate curiosity, thus triggering questions not just about biology of individual organisms and species diversity, but also about agriculture, energy, medicine, public health, and many other issues of critical importance to humanity (Cook et al., 2014). As

---

[11] See http://www.yeastorfanproject.com.
[12] See https://nsf.gov/awardsearch/showAward?AWD_ID=2032158&HistoricalAwards=false.
[13] See https://virtual.imnh.iri.isu.edu.

the foundation for what is known about how life on Earth changes over time and space, biological collections provide windows into the past, providing evidence for how species have evolved and how biological communities have changed through time. During the late 1970s and the 1980s, museums began presenting "glitzy" exhibitions that visitors did not like because there were fewer specimens on display (Hooper-Greenhill, 1994). Today, many universities and natural history museums use public exhibitions demonstrating how their biological collections are unique spaces for interdisciplinary research and educational innovation, providing a place-based window through which to focus on integrating science and discovery (Bakker et al., 2020).

Engaging Lifelong Learners in Citizen Science[14]

Citizen science is another area in which biological collections encourage an interaction between research and education. Citizen science has grown as a way to engage individuals and communities in authentic scientific and inquiry-based activities, increasing public appreciation and support for science and serving as a valuable contributor to advancing scientific research (NASEM, 2018d). Many biological collections, particularly natural history collections, actively pursue projects to include people, many of them without professional training in science, in a wide array of collections-related endeavors. These activities can range from supporting digitization efforts to participating actively in the science as data collectors or lab assistants identifying critical taxonomic features of particular specimens. For example, in 2012, natural history collections professionals partnered with experts in citizen science and data visualization to create Notes from Nature, a "prototype citizen science application" that enabled volunteer members of the public to help digitize specimen labels and field notes (Hill et al., 2012). Notes from Nature is one of many scientific projects on Zooniverse, a popular Internet platform for volunteer-based scientific research. Since 2012, more than 8,200 volunteers have completed more than 1.1 million transcriptions.[15] Similarly, Worldwide Engagement for Digitization Biocollections (WeDigBio),[16] which launched in 2014, is an international citizen science project to create digital data from specimens (Ellwood et al., 2018). Each year WeDigBio hosts a 4-day event during which volunteer members of the public can visit local museums, universities, field stations, marine laboratories, and other organizations to help scientists create specimen data using online platforms such as Notes from Nature (see Figure 3-1).

Some natural history collections also host or participate in programs such as Bumble Bee Watch,[17] a citizen science project to track and conserve bumblebees

[14] Citizen science refers to "people who are not professionally trained in disciplines relevant to a specific project participating in the processes of scientific research, with the intended goal of advancing and using scientific knowledge" (NASEM, 2018d, p. 1).

[15] To help digitize specimen labels and field notes, see https://www.zooniverse.org/organizations/md68135/notes-from-nature.

[16] See https://wedigbio.org.

[17] See https://www.bumblebeewatch.org.

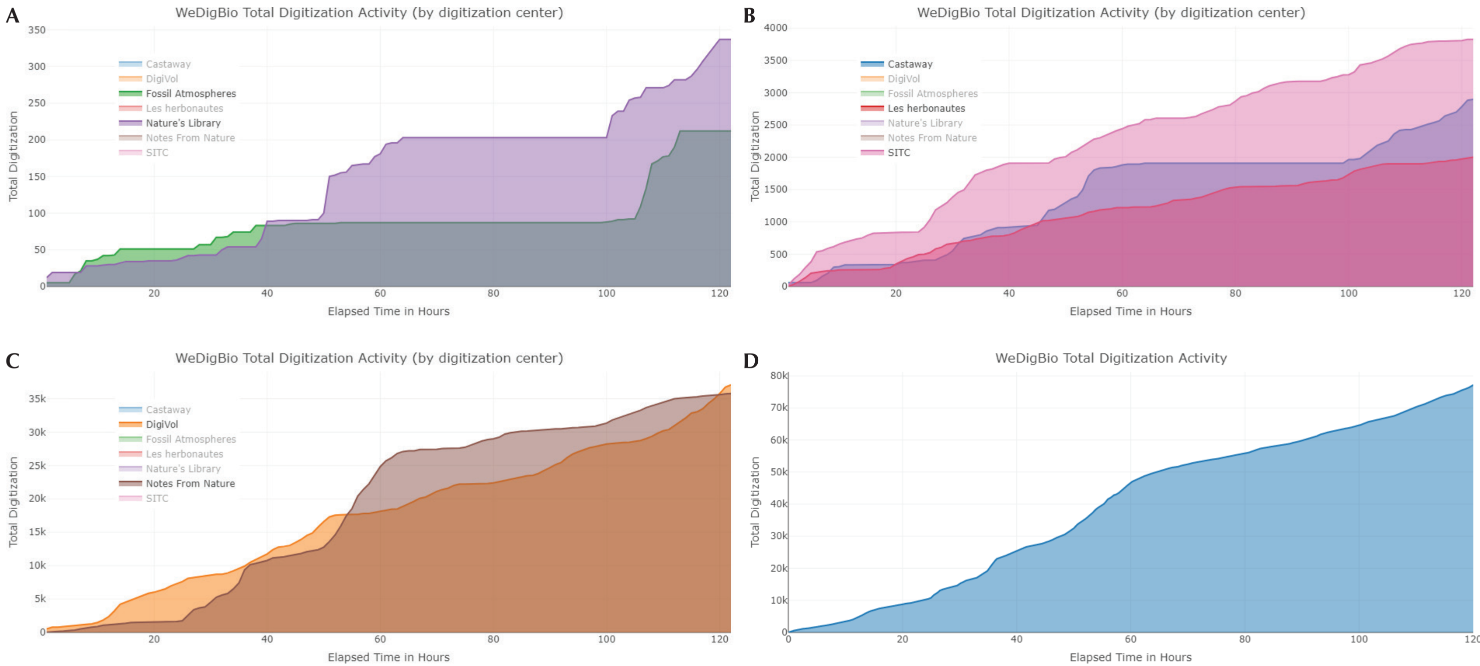


FIGURE 3-1 WeDigBio Total Transcription Activity during the 2019 Annual Event (by transcription center). The 2019 WeDigBio annual event leveraged seven online platforms. This figure shows the total number of digitization activities (e.g., transcriptions) that took place over a 120-hour period. A: Fossil Atmospheres and Nature's Library. B: Castaway, Les Herbonautes (supported by the Muséum National d'Histoire Naturelle, Paris, France), and the Smithsonian Institution's Transcription Center. C: DigiVol (supported by the Australian Museum) and Notes from Nature (part of Zooniverse). Volunteers at DigiVol and Notes from Nature contributed to the greatest number of digitization activities (>35,000 each). Volunteers contributed up to 4,000 activities at each of the other platforms. D: The cumulative number of digitization activities across all seven platforms in the 4-day period was 77,154.
SOURCES: WeDigBio 2019 Dashboard, image courtesy of Austin Mast.

in North America. Bumble Bee Watch engages collections professionals at the Natural History Museum, London, the Montreal Insectarium, and several other scientific institutions to help verify the identities of bumblebees in community-submitted photographs. Funders, collaborators, and experts come from all over the world, and several regional efforts, such as the Maine Bumble Bee Atlas,[18] add further support for this endeavor as well.

BROADENING PARTICIPATION IN SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS

Multiple reports emphasize the value and importance of diversity, equity, and inclusion in STEM disciplines and underscore the need to broaden participation of underrepresented groups, including women and racial and ethnic groups (NAS et al., 2011; NASEM, 2016, 2018c, 2019c; NRC, 2011). "Encouraging greater diversity is not only the right thing to do: it allows scientific organizations to derive an 'innovation dividend' that leads to smarter, more creative teams, hence opening the door to new discoveries" (Nielsen et al., 2017, p. 1740).

STEM education research demonstrates that inquiry-based learning and undergraduate research experiences, such as those provided by some biological collections, improve student understanding of STEM concepts (NASEM, 2017c) and may be important mechanisms to encourage diverse communities to pursue careers or avocations in STEM (Hernandez et al., 2018). For example, the Girls at the Museum Exploring Science project is a collaborative effort between the University of Colorado Boulder, the University of Colorado Boulder Museum of Natural History, and the Boulder Valley School District (14 elementary schools). It is an ongoing 7-week afterschool program, designed exclusively for girls in the fourth and fifth grades from diverse and underrepresented racial and ethnic groups. Creating safe spaces in informal contexts is effective in changing the girls' interests in and attitudes toward science, influencing future education, careers, leisure pursuits, and ways of thinking about what science is and who does it, as well as shaping their personal identities, life trajectories, and social, cultural, and science capital (Archer et al., 2015; McCreedy and Dierking, 2013).

Another example in which natural history collections have been used to broaden participation in STEM is through a Columbia University–based project, Early Engagement in Research: Key to STEM Retention, supported through an NSF INCLUDES Planning Grant.[19] This project enables high school students from communities historically underrepresented in STEM to work on specific Earth and environmental science challenges with college students, science teachers, and research experts. Public land and resource management agencies (New York City Department of Parks & Recreation, U.S. Fish & Wildlife Service, and Department of Agriculture Forest Service) provide access to field and research sites, along with research dissemination opportunities. Research projects involve biological collections and study the consequences of reforestation in the New

[18] See http://mainebumblebeeatlas.umf.maine.edu.
[19] See https://www.nsf.gov/awardsearch/showAward?AWD_ID=1359194&HistoricalAwards=false.

York City ecosystem, providing scientific support for management of invasive and rare species in the region. In addition, iDigBio holds workshops to address broadening participation in the biological sciences with the goal of introducing students, especially those in underserved populations, to museum and biodiversity science careers.[20]

EVALUATING IMPACTS ON FORMAL EDUCATION AND LIFELONG LEARNING

Though biological collections have a rich and long history of being used in educational activities, there is very little documentation about collections' specific impact on student learning in schools or on lifelong learners. For example, it is known that museum experiences, both for schoolchildren and lifelong learners, can result in learning (Falk and Dierking, 2018; Mujtaba et al., 2018). However, the role of biological collections in these museum experiences is implicit, rather than explicit. Many classroom lessons, public exhibitions, and citizen science programs are evaluated, and some, particularly NSF-funded efforts, are even researched, but the value added to such programs by the specific and intentional use of biological collections has yet to be robustly documented[21] or aggregated across projects. Evaluating the impacts of biological collections-based education and lifelong learning endeavors could enable a greater sense of whether and how engaging with biological collections results in better understanding and helps to meet the known learning needs of K–12 students, university students, and members of the public. Evaluation and research could also help to identify the types of learning programs that may be effectively scaled up and used more extensively across the nation for biological collections–based STEM educational activities and other learning endeavors. Although the focus in this section is impact evaluation, there are also evidence-based tools to determine what learners know about the topic or scientific process being proposed for the activity, and strategies to test programmatic goals during a pilot phase, in order to adjust the idea and maximize its impact, once it is implemented (see Box 3-3). Ideally, collaborations will develop among evaluators, education researchers, and biological collections experts, particularly among those employed by the same institution, to select appropriate evaluation tools and develop metrics that provide evidence for the impacts of using biological collections in learning.

Designing, implementing, and expanding the use of collections-based educational programs requires comprehensive planning and dedicated stewardship in order to meet the needs of schools, museums, and other institutions of formal and informal learning. The STEM education research community has many resources to develop and evaluate educational activities and assess learning

[20] See https://www.idigbio.org/content/broadening-participation-biology.

[21] Most educators define "evaluation" and "assessment" differently. Evaluation typically refers to whether and, if so, the degree to which intended goals for a specific education program are achieved and, consequently, whether the program is effective. Assessment refers to measuring changes in an individual's understanding, skills, attitudes, perceptions, beliefs, or other learning-related outcomes.

outcomes (Friedman, 2007; Patton, 2018) (see also Box 3-3). Chapter 2 provides a more in-depth, step-by-step description of best practices for evaluation in the context of documenting the impacts of biological collections on research; many of those principles also apply to education. In brief, the first step is to develop a clear program plan that identifies for whom the learning experience is designed, the goals and objectives for the learning activity or lesson, and why the activity or lesson is important for the intended learners. Being clear about the intended value-added benefits from the start can help biological collections be used in the most effective and strategic manner. Because the primary focus of most biological collections is research, experts in STEM education research,

BOX 3-3
Tools for Developing and Evaluating Science, Technology, Engineering, and Mathematics Education Programs and Learning

K–12

The Next Generation Science Assessment portal describes an evidence-centered design process, tools, and strategies to develop classroom-based science assessments.[a]

Undergraduate and Graduate Education

Community colleges are a critical component of the undergraduate education system as they are widely dispersed around the United States, can quickly adapt to the changing science, technology, engineering, and mathematics (STEM) workforce needs, and reach a broadly diverse group of students (NRC, 2012b). The National Academies report *Indicators for Monitoring Undergraduate STEM Education* published a conceptual model that outlines three primary goals for undergraduate STEM education: (1) increase students' mastery of STEM concepts and skills; (2) strive for equity, diversity, and inclusion; and (3) ensure adequate numbers of STEM professionals (NASEM, 2018c). It lays out ideal indicators and data sources for measuring these goals, many of which are also relevant for graduate education.

Informal Education

The National Research Council report *Learning Science in Informal Environments: People, Places, and Pursuits* (NRC, 2009) outlines the opportunities to be realized with a broader definition of science learning, and ideas for documenting evidence in these areas, including a set of outcomes. The report also outlines six strands of learning that can guide the development of effective educational programs and assessment:

Strand 1: Experience excitement, interest, and motivation to learn about phenomena in the natural and physical world.

professional evaluators, and educators are essential collaborators and partners as strategic educational goals and program plans are developed. Such partnerships can be more feasible when the potential collaborators work in the same institution. Before the program plan is implemented, a strategy to "measure" its impact through some form of evaluation is needed. It is important to note that evaluation is a set of processes and tools to document the outcomes and accomplishments. Metrics will vary depending on the goals of an educational effort and on whether impacts are being measured with K–12 students, undergraduate and graduate students, lifelong learners, volunteers, or citizen scientists who interact with collections-based programs or exhibitions.

Strand 2: Come to generate, understand, remember, and use concepts, explanations, arguments, models, and facts related to science.

Strand 3: Manipulate, test, explore, predict, question, observe, and make sense of the natural and physical world.

Strand 4: Reflect on science as a way of knowing; on processes, concepts, and institutions of science; and on their own process of learning about phenomena.

Strand 5: Participate in scientific activities and learning practices with others, using scientific language and tools.

Strand 6: Think about themselves as science learners, and develop an identity as someone who knows about, uses, and sometimes contributes to science.

Citizen Science

The National Academies report *Learning Through Citizen Science: Enhancing Opportunities by Design* (NASEM, 2018d) describes how citizen science projects can support a variety of learning outcomes. Some of these outcomes, such as developing motivation and learning new scientific skills, are relatively common within the activities and practices used across all citizen science projects. Others, such as encouraging the development of scientific reasoning, come only with significant supports and scaffolding. However, there are few investigations into the unique learning opportunities associated with citizen science, though the work around identity development in citizen science heads in this direction (Ballard et al., 2018). Because citizen science invites nonscientists into science, it provides an opportunity to welcome and explore differing cultural perspectives and how they may enrich science learning *and* science overall. This has the potential to shed light on the persistent historical underrepresentation and underparticipation of many communities and their members in science, insights that are likely to be useful well beyond citizen science.

[a] See http://nextgenscienceassessment.org.

CONCLUSION

There is a long-standing tradition of biological collections contributing to educational endeavors. Many of those endeavors in formal and informal education align with evidence-based principles known to stimulate interest and excitement in learning, increase scientific knowledge, and improve the understanding and use of scientific practices and tools. These educational endeavors are rich in diversity and depth, and constitute a unique and important contribution to the nation's efforts to promote lifelong learning in STEM. As the volume and diversity of digital biological collections data expand, the educational opportunities in data science will also expand to complement disciplinary and transdisciplinary learning. Collaboration with experts in educational research, evaluation, and assessment will help to refine biological collections-based educational objectives and programs, determine the impact of those programs on learning, and perhaps help to identify a set of approaches or programs to implement at a national scale.

4

Building and Maintaining a Robust Infrastructure

The health of biological collections—and, ultimately, of the scientific research that relies on them—is dependent on the underlying infrastructure that assembles, maintains, and provides access to these collections. That infrastructure includes not only the physical space and equipment used to house and maintain the specimens in a collection, but also their accompanying data and the procedures governing their care. It includes the technologies to produce digital data and the cyberinfrastructure to store, analyze, and aggregate data with those of other collections through online portals (see Chapter 5). Finally, biological collections infrastructure includes the trained staff, students, and volunteers who acquire, curate, manage, and ensure the quality of specimens and their data, and coordinate their scientific and educational uses. Such infrastructure can be expensive and time consuming to maintain, but the value that biological collections provide to the scientific research and education communities more than justifies these expenditures. For example, an analysis of biological resource centers that collect, certify, and distribute living organisms demonstrated that these institutions amplify the cumulative impact of individual research discoveries and thereby significantly increase the pace and reduce the cost of research (Furman and Stern, 2011).

This chapter focuses on the physical infrastructure challenges of ensuring that biological collections remain available and viable for research and educational use. It also touches on an important aspect of the biological collections infrastructure—the mechanisms that ensure that the extended research and the broader education communities have convenient and effective access to the biological specimens maintained in these collections.

THE PROMISE OF BIOLOGICAL COLLECTIONS INFRASTRUCTURE

It is easy to overlook the importance of infrastructure. When everything is functioning smoothly, infrastructure—whether it is the facilities of a university, the computers and transmission devices underlying the Internet, or the air traffic control system responsible for air travel—tends to be taken for granted.

The same is true of the nation's system of biological collections. When collections are discussed, it is generally in terms of their physical, digital, and intellectual assets and resources used by researchers and others to answer questions about past, present, and future life on Earth. But those resources are available only because of the nation's biological collections infrastructure, which not only maintains the specimens and associated biological materials and data, but also supports the means in which they are widely shared and distributed.

The nation's biological collections have a dual nature similar to that of biological field stations and marine laboratories, which are both individual entities and "collective elements of the nation's broader scientific infrastructure" (NRC, 2014b, p. 45). As individual research repositories, each biological collection serves the institution in which it is housed and also serves the broader scientific community. Individual biological collections vary in nature from small, project-based collections with relatively simple infrastructure needs to large repositories of diverse living, fossil, and preserved specimens and their associated data with complex, sophisticated, and ongoing infrastructure needs.

Biological collections can also be thought of as a collective system that is a vital component of the nation's scientific infrastructure. This distributed system is somewhat analogous to the National Radio Astronomy Observatory (NRAO),[1] a dispersed set of telescopes that provide resources to astronomy researchers worldwide, as well as to formal and informal educational programs. The capability of distributed biological collections to serve as a collective national resource depends on ongoing digitization efforts and a cyberinfrastructure that allows them to link and integrate their digital data (see Chapter 5). One current difference from the NRAO is that biological collections are managed independently, with each collection in the network largely setting its own strategic plan and being responsible for its own mission, management, and funding.

The specific physical infrastructure needs of biological collections vary according to the types of specimens they contain (e.g., size, number, taxonomy, and biosafety level), the maintenance requirements of the specimens (e.g., wet, dry, refrigerated, or frozen), and the intended scientific and educational objectives (see Figure 4-1). The requirements for cryopreserved (frozen) biological collections, for example, are particularly stringent because the specimens and biological material lose their viability or integrity if they thaw. Such collections are often stored in freezers kept at –80ºC or in cryogenic storage drawers using liquid nitrogen at –190ºC, both of which require constant monitoring and backup generators, particularly for specimens without duplicates housed at another location. At a minimum, all biological collections require a secure facility with the necessary equipment and controls to maintain lighting, temperature, humidity, airflow, and other environmental conditions at the levels required to maintain the specimens and prevent contamination and degradation. Many organisms are represented by a variety of collection types that may require different preservation methods, storage conditions, and locations (sometimes even involving multiple institutions). For example, in addition to herbarium specimens, plants may be represented by separate wood or seed collections, cell or callus cultures, plant genes in bacterial plasmids, frozen or silica-dried leaf tissue, and whole plants in fields, greenhouses, or growth chambers. Mammal and bird collections can include live animals, skin, and skeleton (or fluid-preserved) voucher preparations, frozen tissues, cell cultures, embryos, sperm, karyotypes, diverse sets of endo- and ectoparasites, and more (Galbreath et al., 2019). Ichthyology and herpetology collections contain predominantly ethanol-preserved wet specimens,

[1] See http://public.nrao.edu/about.

FIGURE 4-1 Different types of specimen storage. (A) Dry storage: fossil shells in drawers at the University of Colorado Boulder Museum of Natural History. (B) Cryogenic storage: microbial strains at the American Type Culture Collection. (C) Liquid storage: specimens in jars at the Florida Museum of Natural History. (D) Greenhouse at the Arabidopsis Biological Resource Center at The Ohio State University.

but also maintain cleared and stained glycerin specimens and skeletal and tissue collections, all requiring different storage conditions.

Even organisms that appear superficially similar, such as different types of microalgae, may require different types of infrastructure to maintain them as biological collections (see Box 4-1). Because of the wide variety of biological collections, there are many publications that describe specimen-specific infrastructure requirements and baseline standards (see Box 4-2). The basic physical infrastructure requirements also involve a variety of materials, tools, technologies, and other resources necessary to maintain and curate collections. Examples include compactors, digitization infrastructure (cameras, lighting, scanners, printers, etc.), backup generators, safety requirements (e.g., for ethanol or cryogenic collection storage), media and reagents to preserve or promote growth, sensors, and alarms to monitor and raise alerts about unauthorized access or

BOX 4-1
Infrastructure and Maintenance of Microalgae Cultures

Photo A: Algal culture room, courtesy of the Culture Collection of Algae at The University of Texas at Austin; Photo B: Courtesy of the National Center for Marine Algae and Microbiota

Microalgae are single-celled photosynthetic organisms that live in a wide range of aquatic and semi-aquatic habitats including lakes, rivers, oceans, snow, and damp soils. Collections of microalgae are used for a variety of research and commercial applications, such as for biofuel production, drug and nutrient development, and cosmetics (Khan et al., 2018). Some species of microalgae can be cryopreserved (frozen), but others must be maintained as live cultures. The UTEX Culture Collection of Algae at The University of Texas at Austin maintains algal live cultures on agar and in liquid, while some are cryopreserved. The National Center for Marine Algae and Microbiota (NCMA), located at the Bigelow Laboratory for Ocean Sciences in East Boothbay, Maine, is an example of two private culture collections that were developed into a national resource center to meet the needs of the research community. The NCMA maintains the largest and most diverse collection of publicly available marine microalgal strains.

fluctuations or unsafe environmental conditions, and tools and technologies to authenticate accessioned material and periodically assess the condition or determine the genetic identity of specimens.

An important feature of biological collections is that they—like many culturally and historically important collections—continue to grow. Specimens are added to biological collections through three main mechanisms: (1) field collecting of specimens in previously unexplored ecosystems and resurveying previously sampled ecosystems; (2) generating new, living genetically modified research organisms; and (3) the acquisition of specimens or entire collections by gift, donation, exchange, or purchase. During the 19th century, many of the largest and most ambitious biological collections grew through specific national or international research mandates to catalog all species of a given region, taxon, or clade. Today, many biological collections grow principally as a product of individual research projects or an individual institution's priorities.

The potential ramifications of neglecting the nation's biological collections infrastructure are wide-ranging, with severe consequences for innovations in biotechnology, medicine, agriculture, energy, and many other sectors built on life science research (Flattau et al., 2007; McCluskey, 2017; Sigwart, 2018). Neglecting infrastructure could also affect research, public services, and private businesses that rely on accurate taxonomic identification, such as forensics, the study of disease outbreaks (human, wildlife, and agricultural), border protection, and the control of invasive species (Cook et al., 2020; McLean et al., 2016). In addition, most natural history collections are non-renewable scientific resources—they cannot be replaced. The loss of individual specimens or entire collections creates unfillable gaps in the knowledge of present and past life on Earth. Institutions that do not provide adequate infrastructure for their biological collections hamper their own missions to advance science and technology, build a highly skilled workforce, and educate the next generation of global citizens.

CHALLENGES

Maintaining a healthy physical infrastructure involves a variety of interrelated challenges. Perhaps the most obvious challenge involving specimens and data is that they need to be preserved indefinitely, beginning with their initial accession and continuing with long-term maintenance for both anticipated and unanticipated uses. Accordingly, the quality of the specimens needs to be carefully and constantly maintained to ensure that findings from past research can validly and reliably be compared with results in any number of future research investigations. These challenges are exacerbated by the fact that healthy collections are continually expanding through the acquisition of new material, which requires a steady increase in physical capacity. Finally, making specimens and data available to researchers and other users, including educators, students, and businesses, is important in maximizing the usefulness and of impact infrastructure considerations for the nation's biological collections. The following sections describe these challenges in more detail.

Collections Require Ongoing Preventive Conservation

Without active and ongoing preventive conservation,[2] natural history specimens will degrade over time and become less useful for research and education. Fluid-preserved specimens will eventually dry out if not stored in appropriate containers and resupplied with the appropriate liquids, cryopreserved tissues will decay if freezers are not maintained and kept at desired temperatures, dried collections can fall victim to insects and mold, and fossils are subject to Byne's (Cavallari et al., 2014; Shelton, 2008) and pyrite diseases (Cavallari et al., 2014; Hall, 1998, Larkin, 2011). Responding to the requirement that collections be

[2] Preventive conservation is defined as actions taken to minimize or slow the rate of deterioration and prevent damage; it includes activities such as risk assessment, the development and implementation of guidelines for continuing use and care, ensuring appropriate environmental conditions for storage and exhibition, and instituting proper procedures for handling, packing, transport, and use (SPNHC, 1994).

BOX 4-2
Select Publications About Requirements and Standards for Biological Collections Infrastructure

Preventive Conservation: Collection Storage Elkin and Norris (2019)	A comprehensive reference for a risk-management approach for all types of collections including fine arts, libraries, and biological collections. It discusses planning and assessment, building design and facilities management, and storage furniture and specimen housing.
The Biological Resources of Model Microorganisms Jarrett and McCluskey (2019)	A comprehensive reference on the living stock collections of 14 different model organisms. It provides the history of each model organism, how the organisms are being used in scientific research, and the particular requirements and best practices to obtain, maintain, preserve, characterize, and distribute the organisms.
Herbarium Practices and Ethics III Rabeler et al. (2019)	A scientific publication that provides recommendations and key considerations for the infrastructure, operation, and services of herbarium collections, including digitization and virtual capabilities. It is the third update of a 1958 publication.
ISO 20387:2018: Biotechnology—Biobanking—General Requirements for Biobanking International Organization for Standardization (ISO, 2018)	International standards that define the basic requirements for the competence, impartiality, and consistent operation of biobanks.
Best Practices: Recommendations for Repositories, Fourth Edition International Society for Biological and Environmental Repositories (2018)	A comprehensive reference on the technical and managerial requirements for biological repositories, including storage and processing equipment, information management systems, business planning, and specimen collection and access, among other critical dimensions. Campbell et al. (2018) provides a brief, accessible guide to new and revised details included in the fourth edition volume.

Health and Safety for Museum Professionals Hawks et al. (2010)	A three-part publication that provides guidance on facilities management, infrastructure, and functions of museum staff to ensure a safe and hazard-free collection. Some of the issues addressed include fire protection; occupational and hazardous waste management; chemical, physical, electrical, and radiation hazards; and energy salvage, field work, conservation, and restoration. This publication is the result of a collaboration between the American Institute for Collaboration and the Society for the Preservation of Natural History Collections (SPNHC).
Best Practice Guidelines for Biological Resource Centers Organisation for Economic Co-operation and Development (OECD, 2007)	International guidelines that address the full portfolio of infrastructure and management needs to maintain the quality and services provided by biological resource centers, including potential approaches to national certification. The guidelines resulted from discussions among the OECD member countries, key partner countries, and the scientific community to serve as a target for quality management of living stock collections.
Storage of Natural History Collections: Ideas and Practical Solutions Rose and de Torres (2002)	A comprehensive compendium of 113 articles on the practical applications of storage systems for everything from vertebrate teeth to ethnic costumes to large fossils. Each article was written and reviewed by professionals in the fields of conservation and collections management.
Managing the Modern Herbarium Metsger and Byers (1999)	A comprehensive reference on a wide variety of collection care and management topics, including environmental controls, pest management, paper conservation, adhesives, destructive sampling, and case studies on moving a herbarium to new quarters. It is the result of a collaboration between the Royal Ontario Museum and SPNHC.

viable and pure, living collections also address these issues through quality control processes as described below. Providing ongoing funding for the active care of collections—as well as for the accessioning of new specimens into collections—is a challenge for an institution, especially one charged with the maintenance of many types of scientific research infrastructure.

In addition, many biological collections are located in environments that are prone to disaster—natural and human-caused. For example, an attempt to assess risks to herbaria found that about half of all herbaria have at least three risk factors, one of which relates to their location in areas prone to flooding, earthquakes, severe weather (hurricanes, typhoons, etc.), or social unrest, and the other risk factors relate to insufficient staffing and limited utility to modern research because of the collections' inaccessibility (Thiers et al., 2018). However, even in relatively safe locations, inadequate infrastructure enhances vulnerabilities to natural disasters and theft (Araujo, 2019; University of Vermont, 2017). Although it may be possible to recover from damage to facilities and equipment, many natural history specimens, including fossils and specimens collected in the past, contain baseline knowledge for historic environmental conditions and prior research that cannot be replaced. The coronavirus disease 2019 pandemic poses an additional threat. Up to one-third of all museums in the United States may permanently close due to financial losses during the pandemic, potentially leading to the loss or relocation of millions of natural history specimens and fossils.[3] In addition, some living collections cannot be cryopreserved or lyophilized, such as some microalgae or *Drosophila* species, and require labor-intensive procedures to maintain. When fewer people are allowed to enter the facility during a pandemic, maintenance of these stocks can be impacted and if their transfer is delayed, stocks may lose viability.

Sometimes the only solution to failing infrastructure is to transfer specimens wholly or in part to a more stable situation. Usually, the collections transferred are small, although in 2018, the University of Louisiana Monroe moved nearly 6 million specimens from the herbarium and fish collections to the Botanical Research Institute of Texas and Louisiana State University (the herbarium collection) and to the Tulane University Biodiversity Research Institute (the fish collection). Many collections have been saved from decline or outright destruction when rescued by another institution that was willing to accept responsibility for their care. The United States Culture Collection Network published a survey of rescued living microbial collections, including resources at the *E. coli* Genetic Stock Center, the Fungal Genetics Stock Center, the Phaff Yeast Culture Collection, and the Culture Collection of Algae at The University of Texas at Austin, and described some of the scientific discoveries made with the rescued specimens since then (Boundy-Mills et al., 2019). This effort demonstrated the value of having established capacity to ensure that important collections survive when there is insufficient financial support or when senior staff retire or change institutions. However, transferring collections to a new institution may potentially have negative consequences. Such transfers increase infrastructure, financial,

[3] See https://www.youtube.com/watch?v=48rKE129ME4.

and regulatory requirements at the new institution, break links to historical knowledge about the collection, and can remove the specimens farther from the region where the specimens were collected and, potentially, from the primary users of those collections.

Living Stock Collections Require Consistent Quality Control

The quality of a living stock collection is a major determinant of whether its specimens can be used for research and of the type of research for which they are most suitable. Specimen quality is critical for ensuring the reproducibility and replicability of research results and reflects on the credibility of collections and their institutions.

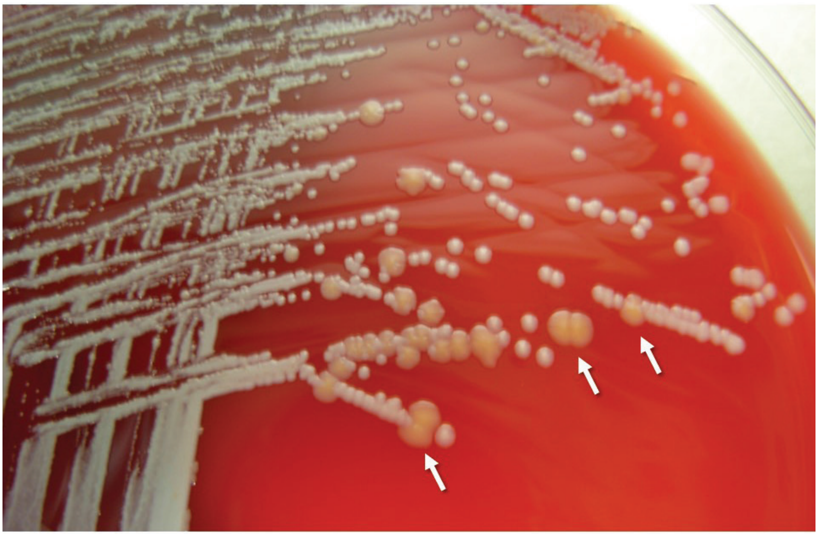
Quality control, which is similar to preventive conservation for natural history specimens, is the process through which collections personnel seek to ensure that the quality of specimens and reference materials, such as cell cultures, are standardized and maintained. Customers of living stock collections expect that the material they receive will be properly identified as to the species, will possess the expected genetic markers, and will be viable and pure. For these reasons, many living stock collections have staff dedicated to *quality assurance*. Quality assurance documents and demonstrates control over the quality control processes. Quality assurance facilitates and organizes historical information about the origin and handling of the material and also preserves the traceability of the material.

Quality control and quality assurance require performing standardized tests for authentication, sample characterization, replenishment, and long-term stability. Such standardization is based on experience previously gained and includes predetermined ideal ways to identify suitable growth, storage conditions, and protocols to characterize and define the biological materials. These efforts extend not only to handling the materials, but also to shipping the materials to users and receiving incoming materials.

Every living collection has taxon-specific minimum categories of quality benchmarks. Jarrett and McCluskey (2019) describe some of the quality considerations for 14 different living model organisms along with descriptions of the facilities and procedures necessary to maintain them. Box 4-3 provides an example of typical minimum categories of quality benchmarks for living microbial collections.

There are four key challenges to maintaining quality control of living collections. First, best practices are neither standardized across the living stock collections community nor updated as new regulations and technologies become available. Second, some living collections, such as those of bacteria, yeast, fungi, and other microbes, contain specimens isolated so long ago that they need to be re-identified using current taxonomy and technologies. Third, it is often difficult to confirm genetic markers in materials received from the research community because many living stock collections lack access to specialized personnel, reagents, and equipment for genotyping. Fourth, the equipment and infrastructure to cryopreserve

BOX 4-3
Common Benchmark Categories That Define Quality Control for Living Microbial Collections



Culture of Staphylococcus aureus *grown on blood agar. During growth, this item showed two colony types, suggesting the presence of a contaminant. The arrow shows bigger and pigmented colonies that, after further analysis, were confirmed to be the contaminant. The smaller white colonies were colony-purified for distribution. SOURCE: Biodefense and Emerging Infections Research Resources Repository.*

living collections are expensive. Cryopreservation using liquid nitrogen tanks is an effective, but costly, approach to ensuring the longer-term viability of many types of cells and tissues. Many living collections opt for mechanical freezers, which are more affordable but result in a reduction in long-term viability. For some organisms, lyophilization (freeze-drying) may be used, rendering the material stable for long periods of time at room temperature.

Collections Need Room to Grow

There is a pressing need for the strategic expansion of the nation's set of biological collections to ensure they adequately represent the diverse array of Earth's biota across space and time. The continual growth of biological collections is essential for tracking ongoing global change, especially because the planet's habitats and physical environments are now rapidly shifting. Given the

Viability. Living collections need to define protocols to confirm that their microbes pass quality control after amplification, preservation, and shipping to the user. Some collections often perform viability testing at intervals during storage. Vertebrate facilities use best practices in alignment with the Animal Welfare Act (https://www.nal.usda.gov/awic/animal-welfare-act), which may include testing for pathogens and health status prior to distribution.

Identification. Confirmation of the specimen identity down to the genus and species level (if known) is done for each lot. Accurate identification and characterization of the material are crucial for compliance with regulations related to restricted agents. For example, organisms that can be weaponized are highly regulated and controlled. These collections require a high investment in infrastructure, which explains why these organisms are handled by very few biological collections.

Purity. This category applies to microorganisms as well as larger organisms. Microbes are confirmed pure by standard macroscopic and microscopic techniques as well as molecular assays such as nucleic acid sequencing. Collections of larger organisms such as plant germplasm (http://fps.ucdavis.edu), zebrafish (zebrafish.org), and *Xenopus* (https://www.mbl.edu/xenopus) are subjected to a sanitation and/or quarantine process to avoid contamination of the facility.

Strain characterization. Some microbiological collections have defined processes to follow, depending on the collection, the organism, the intended use of the material, etc. Unique characteristics of the microorganism need to be confirmed. For example, bacterial isolates for research focused on antibiotic resistance might require confirmation of their antibiotic resistance patterns.

tremendous anthropogenic changes now under way, sampling and archiving the baselines of the presence and distribution of organisms will support future scientists in their efforts to understand changes in biodiversity and organisms' responses to global change. Likewise, the expansion of living stock collections, including both new types of genetic stocks and new types of products from existing specimens (e.g., tissues, clone libraries, or purified genomic DNA), is essential for many services and areas of research and development, including synthetic biology, microbiome analysis, bioterrorism, and developing crops and livestock able to thrive in an altered climate. Regardless of the reasons, the growth of biological collections requires strategic thinking about infrastructure from multiple angles—the capability to expand space, the development of tools and technologies that help reduce space required for specimen maintenance and storage, and the more effective use of existing space.

A collection that has stopped growing is often seen by others in the community as being inactive and thus may be overlooked as a research resource. Some biological collections do not have general growth, or even strategic growth, as part of their mandate. Those biological collections that include growth in their mandate may vary widely in the degree to which they pursue growth. For example, the ornithology collection of the Burke Museum of Natural History and Culture at the University of Washington in Seattle has pursued an aggressive policy of growth since its founding in the 1970s and is now, after five decades of sustained growth, one of the premier ornithology collections in the world. Similarly, the insect and other collections generated by the National Ecological Observatory Network (NEON)[4] are being housed at Arizona State University, which has allocated approximately 10,000 square feet for this purpose (per personal communication, Nico Franz, curator, Hasbrouck Insect Collection, Arizona State University, November 2019). Many living microbial collections include growth within their mission, in part because many scientific journals require that microbial strains used in publicly funded research be made accessible to the research community for future study. Unfortunately, many living collections lack the capacity to accession a high volume of material from publicly funded research, even if collecting such material is within their mission. Addgene,[5] a nonprofit global plasmid repository, is an example of an independent entity that accepts, archives, and distributes thousands of plasmids, viruses, and other materials cited in research publications. However, the Addgene model has not yet been applied to engineered or constructed living strains used in research.

Typically, collections growth is the result of funding for specific research projects that have a specific focus on a particular taxon or on developing a new type of research organism (e.g., an organism with specific genetic modifications). In other words, growth is typically not the result of a coordinated collecting strategy. As a result, many collections have well-known biases in terms of species, sex, size, or the geographic distribution of specimens, and correcting such biases can be an important motivation for continued growth. Often, the growth of biological collections creates tension between the resources needed to curate and maintain existing collections and the resources needed to house and manage incoming biological material. Additionally, growing biological research collections may compete for space with other institutional functions (e.g., classrooms, research laboratories, and athletics), some of which may be deemed more relevant for immediate revenue generation or the mission of the larger institution.

In general, the infrastructure funding programs of the National Science Foundation's (NSF's) Division of Biological Infrastructure (DBI) do not include provisions to ameliorate the demands that collections growth places on biological infrastructure. DBI's Collections in Support of Biological Research program explicitly excludes what it deems as "normal" growth, even though there are no clear metrics by which normal growth is determined. NSF provides support for

[4] See https://www.neonscience.org.

[5] See https://www.addgene.org.

growth in only two situations: when there is an urgent need for an institution to subsume an orphan (abandoned) biological collection, or for new collections produced from national and international initiatives such as NEON. DBI's Sustained Availability of Biological Infrastructure program, established in 2019, only provides support to prevent the loss of "mature" physical infrastructure and cyberinfrastructure. Notably, NSF does not require research proposals that involve collecting or generating new specimens to include support for collections maintenance and growth. All research proposals are required to include a data management plan[6] to describe how research results, including data from specimen-based work, will be disseminated and shared. However, there is not yet a requirement for a *specimen management plan* to describe how specimens and their associated data will be curated, digitized, and cared for over the long term for an established biological collection. Additional discussion about the need for a specimen management plan, including the management of the digital data associated with specimens, and requirements for an accompanying budget to support the management plan is offered in Chapters 5 and 7, respectively.

It is possible that the growth of biological collections is not recognized as a pressing problem and so it has not traditionally been a primary criterion for NSF to grant infrastructure funding. Yet, many improvements in infrastructure, including increased space, compactors, and robotic access to specimens and other facilities, can ameliorate the challenges of collections growth. The lower priority placed on growth may have stemmed from the fact that, over the past few decades, collections growth has slowed for many institutions (Malaney and Cook, 2018). The reasons for this slowdown are varied, including a lack of physical space for new collections, an increased reliance on project-based collecting, increased difficulty obtaining permits and navigating the increasingly complex legal issues surrounding biological collecting, the perception of leadership at host institutions that collecting and the fieldwork associated with it are not valuable, and changing societal norms surrounding biological collecting (Antonelli et al., 2018; Bakker et al., 2020; Wallace et al., 2013).

Biological Collections Need to Be Accessed

Open science[7] is a major, global trend that is changing the culture and practice of science. Open science facilitates the exchange of not only biological materials, but also of ideas, data, and other resources such as databases, journal publications, and analytical software (Becker et al., 2019). In the context of open science there are three interrelated challenges facing the accessibility of biological collections: (1) discoverability, (2) physical access to specimens, and (3) access to digital specimen data (see Chapter 5).

[6] See https://www.nsf.gov/pubs/policydocs/pappg20_1/pappg_2.jsp#IIC2j.

[7] Open science is transparent and accessible knowledge that is shared and developed through collaborative networks (Vicente-Sáenz and Martínez-Fuentes, 2018). See https://doi.org/10.1016/j.jbusres.2017.12.043.

Discoverability

The lack of a registry or catalog for all biological collections in the United States is an impediment to open science. Some well-curated catalogs exist for particular types of biological collections. For example, the World Federation for Culture Collections (WFCC) maintains both Culture Collections Information Worldwide,[8] a registry of more than 800 culture collections, and the Global Catalogue of Microorganisms,[9] a public online database of bacteria, fungi, and archaea held in more than 130 collections across 49 countries (Wu et al., 2013). However, because many collections do not have an online catalog of their holdings, users need a catalog or registry that provides collection descriptions in order to find specimens relevant for their research; the lack of such a registry or catalog complicates this sort of discovery.

Access to Physical Specimens

Access to specimens and their related data is of crucial importance to many areas of research and innovation, education, and public engagement in science. However, some biological collections lack adequate space, staff, and research tools for users to study specimens on-site. For both natural history and living stock collections, specimens or the associated biological materials are often shipped to users rather than accessed at the collection facility itself, although this may not be possible if a large amount of material is requested or if the specimens are too fragile or bulky to be shipped. Thanks to national or even worldwide networking, some biological collections can facilitate access to samples that are not stored in their own facilities.

The management of living material requires specific infrastructure such as a laboratory, a greenhouse, or a vivarium as well as the relevant training and expertise. Some living stock collections are only accessible to registered or qualified users. Direct access is usually restricted when specimens represent endangered species or if the materials pose biosafety or biosecurity risks, in which case the user needs to be prequalified to handle the material appropriately in order to minimize these risks (see Box 4-4). Certain microbes and derivatives could potentially be misused and are under strict regulations and controls. The few collections that manage these agents are also under strict control and regulations. Nonetheless, access to these collections is essential in providing support to the scientific community to develop effective countermeasures and control strategies.

Meeting the Needs of a Dynamic Scientific Enterprise

The culture and practices in the scientific enterprise are complex and shifting in several ways that have important implications for infrastructure. First, research

[8] See www.wfcc.info.
[9] See gcm.wfcc.info.

institutions and funders increasingly emphasize and value convergence of scientific disciplines in order to facilitate collaborative, transdisciplinary research and innovation, particularly to address pressing challenges such as antimicrobial resistance, food security, biodiversity loss, and the independent and sustainable production of energy (Jahn et al., 2012; NRC, 2014a). Research infrastructure that promotes convergence and weakens disciplinary silos typically requires physical space that is easy to access but is outside the domain of any single disciplinary department. A hub-like location, for example, the University of Idaho's Integrated Research and Innovation Center,[10] has a variety of design elements that encourage scientists, students, and others to interact formally and informally. Second, institutions that provide formal and informal education programs increasingly support experiential learning in science, technology, engineering, and mathematics (STEM) (Monfils et al., 2017; NASEM, 2017c). STEM education research provides robust evidence that active learning increases interest in and retention of science (NRC, 2015b) (see Chapter 3), thereby making it possible to expand the diversity of the next generation of thinkers who will address ongoing and future challenges facing the planet and human health. In addition, there are growing efforts to cultivate a culture of entrepreneurship and an increasing demand outside of academia for STEM-skilled, workforce-ready graduates. As a result, many colleges and universities are designing (or redesigning) facilities, including research laboratories, classrooms, maker spaces,[11] and informal public gathering spaces that support more immersive transdisciplinary research and experiential learning environments for scientists, students, and learners of all types (e.g., Be a Maker program[12] and the Learning Spaces Collaboratory[13]). The Beaty Biodiversity Museum is a successful example of how a natural history collection might effectively integrate research and educational spaces (see Box 4-5). However, building or renovating space to display collections and create immersive and "hands-on" learning opportunities is financially challenging, particularly for smaller biological collections or those that house sensitive materials. In addition, large-scale infrastructure endeavors to build collaborative, transdisciplinary research and learning environments and the ongoing efforts to address infrastructure needs of biological collections are largely disconnected from one another.

THE WAY FORWARD

Given the challenges described in the previous section, it is clear that new approaches will be required to maintain and improve the value and effectiveness of the nation's biological collections. Growing demands on biological collections will require some fundamental changes to the infrastructure supporting

[10] See https://www.uidaho.edu/research/entities/iric.

[11] Providing the space and the materials for project-based, independent, hands-on experience for students.

[12] See https://beam.unc.edu.

[13] See https://www.pkallsc.org.

BOX 4-4
Infrastructure for Biosafety and Biosecurity of Living Collections

Physical infrastructure–related challenges faced by living biological collections include compliance with increasingly stringent biosafety and biosecurity regulations set by the Department of Agriculture (USDA) and the Centers for Disease Control and Prevention. Some viruses, bacteria, fungi, parasites, protists, or multicellular organisms are pathogenic to humans, animals, or plants. Pathogenicity greatly affects the physical infrastructure required by living collections, for reasons of both biosafety (protecting the safety of the operator handling the organism) and biosecurity (protecting the general public from accidental or intentional release of pathogenic organisms outside the laboratory). For instance, plant pathogens, especially genetically modified plant pathogens, require special use authorization from USDA's Animal and Plant Health Inspection Service and must be shipped under a Plant Protection and Quarantine 526 permit.

Biosafety Level 1 (BSL1) facilities are used for organisms unable to cause disease in humans, animals, or plants. Precautions and infrastructure requirements are minimally restrictive.

Biosafety Level 2 (BSL2) facilities are required for organisms capable of causing disease in humans, animals, or plants but for which the potential diseases are difficult to contract via aerosols. Examples include hepatitis A, B, and C; HIV viruses; and pathogenic strains of *E. coli*, *Staphylococcus*, *Salmonella*, and *Candida*. BSL2 collections must have personnel trained on how to manipulate pathogenic organisms, the personnel must use protective personal equipment, most of the laboratory manipulations should be done within a biological safety

these collections—changes that will grow from new approaches to maintaining them. This section outlines some general strategies that will provide overall improvements in the biological collections physical infrastructure.

Future-Proof the Infrastructure

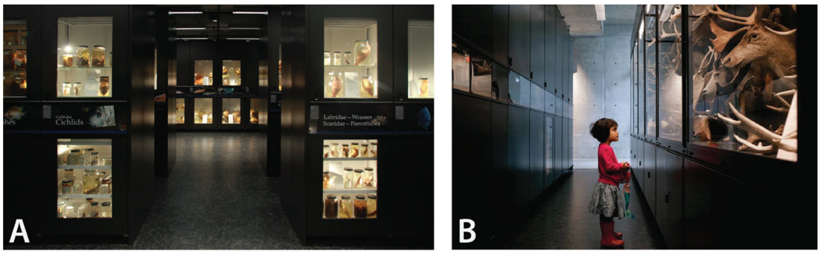
The environment for biological collections is changing rapidly, from new demands being placed on collections by a steady stream of scientific advances to the availability of up-to-date technological capabilities, particularly digital ones, and changes in the ways that scientific research and development are conducted (see Chapter 2). The reevaluation of the collections infrastructure is also motivated by the anticipated increase in the rate of species extinctions (Díaz et al., 2019). It is now incumbent on scientists to approach existing and new specimens, especially those from endangered taxa and threatened biomes, as if it is their last opportunity to do so, because it soon may be. Maximizing cabinet, and the laboratory must have safety protocols in place for decontamination and routine operations.

Biosafety Level 3 (BSL3) facilities are required for organisms that have the ability to infect via aerosols (e.g., *Mycobacterium tuberculosis*), posing a severe threat to laboratory personnel. All BSL2 requirements are followed plus more stringent control on access to the laboratory. Workers require extensive training and certification to work in a BSL3 laboratory. Personnel are under medical surveillance, and respirators or facemasks are required. A hands-free sink and eyewash station must be available near the exit. To minimize the risk of releasing infectious aerosols, floors, walls, and ceilings must be sealed, the laboratory must have negative airflow, the air needs to be filter-sterilized prior to leaving the facility, and the facility must have two sets of self-closing and locking doors. A biosafety manual details all laboratory operations in compliance with all safety requirements, and all work and quantities of materials manipulated must be documented to ensure biosafety. The laboratory must be designed so that it can easily be decontaminated.

Biosafety Level 4 (BSL4) facilities are required to manipulate microbes that could easily be aerosol-transmitted and cause severe to fatal disease in humans and for which there are no available vaccines or treatments (e.g., severe acute respiratory syndrome coronavirus, Ebola and Marburg viruses). These are highly regulated and controlled. Personnel are highly trained and must be approved and certified. Personnel wear positive pressure suits and follow all the requirements and procedures for a BSL3 laboratory. Only a few labs, such as the U.S. Army Medical Research Institute of Infectious Diseases at Fort Detrick, Maryland, meet requirements to handle BSL4 organisms.

biodiversity information for future study requires redoubled efforts to document and preserve it, which will result in many new collections that will need to be accommodated in the nation's biological collections. New methods of propagating living organisms as well as novel methods for preserving tissues and whole organisms may require changes to the current collections infrastructure as well as new curatorial techniques. More robust methods for storage, and the linkage of additional data gathered about endangered and extinct biota, will result in a better understanding of their life histories, habitat requirements, and interactions with other species, and how they reacted to global change in deep time. Guiding this reassessment of current preservation and documentation methods will be the understanding that future knowledge of many species may rely entirely on the specimens and information held in biological collections. Thus, it will be important to ensure that the nation's biological collections continue to thrive no matter what the future brings.

BOX 4-5
The Biodiversity Research Centre's Integrated Space for Natural History Collections Storage, Research, and Education



A: Public Display of Fish Research Collections at the Beaty Biodiversity Museum. Photo by Derek Tan, Beaty Biodiversity Museum, Biodiversity Research Centre, The University of British Columbia
B: Young girl gazing at trophy case interpreting Victorian collecting. Photo by Jeff Werner

The Beaty Biodiversity Museum is part of The University of British Columbia's Biodiversity Research Centre, which integrates space for its natural history collection with public displays, laboratories for collections-based researchers and curators, and offices for educators with related meeting and support spaces. The Beaty Biodiversity Museum, which opened in 2010, includes rows of stacking cabinets with windows, offering visitors views of the research collections, in addition to some small exhibitions. The research center participates in undergraduate and graduate education programs as well as workforce training in biodiversity research. Museum programming, such as Researchers Revealed (https://www.zoology.ubc.ca/~biodiv/rr), is designed to support visitors' understanding of biological collections and their relationship to biodiversity research.

Strategic Planning for Infrastructure

Strategic planning gives an organization the opportunity to evaluate or refine its core mission, identify stakeholders, set goals, and determine the strategies and resources that are needed to achieve those goals. In particular, such exercises require foresight and collaboration between research and administrative staff in an institution to guide the way in which infrastructure challenges are addressed.

Strategic planning can help identify the financial and other needs of a collection and differentiate the funding needed for ongoing maintenance of the collection from that needed to meet evolving standards, replace aging infrastructure, and accommodate the growth of collections. Initiating the strategic planning process every few years can help identify the potential funding sources for biological collections infrastructure and also identify gaps in funding that will need to be met by other resources during the plan's duration (Parsons and Duke, 2013).

Reflections on the core mission of the collection and its primary and secondary stakeholders will help those in charge of the collection come up with actions to ensure the necessary preventive maintenance and quality control of the specimens, increase the specimens' accessibility, and anticipate future uses. The planning process should also take into consideration the availability and training needs of collections leadership and staff (see Chapter 6).

Many collections already engage in regular strategic planning exercises. For the past several years the Society of Herbarium Curators and Integrated Digitized Biocollections (iDigBio) have sponsored a month-long online course entitled Strategic Planning for Herbaria, which trains representatives from up to 10 herbaria to develop succinct strategic plans that include a vision, mission, strategies, and objectives; strengths, weaknesses, opportunities, and threats analysis; sustainability; and assessment and evaluation.[14] Because strategic planning is a common practice for research institutions and universities, it is critical for a biological collection's strategic plan to be part of the plan for the larger institution or to at least be closely aligned to the vision, mission, and goals for the larger institution. Developing an individualized strategic plan requires time, training, and input. Biological collections that do not have the resources to develop a plan could be helped if other collections make their strategic plans publicly available. Sharing strategies to achieve a goal is common practice for federal research institutions and is required or promoted by federal funding agencies and certain universities; such examples could inspire and be used by the broader collection community (North Carolina State University 2017–2022 strategic plan from the Department of Biological Sciences;[15] 2015–2020 strategic plan for the Virginia Institute of Marine Science[16]). Involving an advisory board of experts to help develop and implement the plan could be another way to benefit from the expertise of other biological collections personnel.

Emergency Preparedness

A disaster preparedness and emergency response plan[17] is considered a core document for natural history collections housed in museums and is required for a natural history collection to be accredited by the American Alliance of Museums. Developing a contingency and disaster recovery plan is also recommended for living stock collections (Parsons and Duke, 2013). Such a plan includes responses to natural, mechanical, biological, and human-caused emergencies and addresses the needs of staff, visitors, structures, and collections. However, a preparedness and emergency response plan by itself is no guarantee of successful response to a disaster; in the chaos of an actual emergency it may not be possible to access computers or files where a plan is stored. A regular review

[14] See https://www.idigbio.org/content/strategic-planning-herbaria-short-course-0.

[15] See https://bio.sciences.ncsu.edu/wp-content/uploads/sites/12/2017/08/Biological-Sciences-Strategic-Plan.pdf.

[16] See https://www.vims.edu/about/leadership_admin/dean/strategic_plan_2015/index.php.

[17] See https://www.aam-us.org/programs/ethics-standards-and-professional-practices/disaster-preparedness-and-emergency-response-plan.

of the plan, perhaps with response drills, will keep the actions and supplies needed to recover and stabilize collections at hand in the active memory of collections personnel and allow those personnel to continually refine their plan. An understanding of the special needs of collections by local emergency response agencies may add to the success of a disaster response. In the case of a fire at The University of Vermont's Pringle Herbarium, water damage from hoses was minimized because the local fire department had recently visited the facility as part of a routine check and provided protection against the heat of the fire with padding set onto the tops of the herbarium cabinets (David S. Barrington, The University of Vermont, Director, Pringle Herbarium, personal communication, 2020).

Duplicate Specimens

Depositing duplicate specimens at different institutions can help ensure that specimens are not lost entirely in the event of a disaster. The deposition of duplicate specimens is an established practice among strains of microorganisms, entomological specimens, and herbaria (Groom et al., 2014; OECD 2007; Rabeler et al., 2019). For living stock collections, the Organisation for Economic Co-operation and Development's (OECD's) Best Practice Guidelines go a step further than recommending a duplicate of a specimen be held in a remote location; the guidelines also recommend that specimens be preserved in two or more formats, such as a cryopreserved specimen, lyophilized specimens, or as living cultures. These practices lessen the chance of losses due to power outages, fire, or other types of disasters. For example, copies of several living stock collections are now cryopreserved at the Department of Agriculture National Laboratory for Genetic Resources Preservation in Fort Collins, Colorado (McCluskey et al., 2016), ensuring that these collections can be recovered following a disaster at the home institution. However, it is important to note that the deposition of duplicate specimens is not a practical or even possible solution for many types of biological collections because of already existing issues with space, funding, staffing, and rarity (e.g., dinosaurs or unique culture collection isolates). Nonetheless, when it is possible to have a remote archive of duplicate specimens, this mitigates the risk of specimen loss.

Establish Shared Standards and Technologies for Living Stock Collections

One way to improve the value of living collections is to have strict and consistent quality standards in place. Such standards can help ensure that resources and data are "fit for purpose"—that is, of the type and quality to meet the specific needs of users (Smith et al., 2014). Many companies and organizations follow International Organization for Standardization (ISO)[18] standards, which provide a way to create the documents that provide requirements, specifications, guidelines, or characteristics to ensure that materials, products, processes, and

[18] See https://www.iso.org/standards.html.

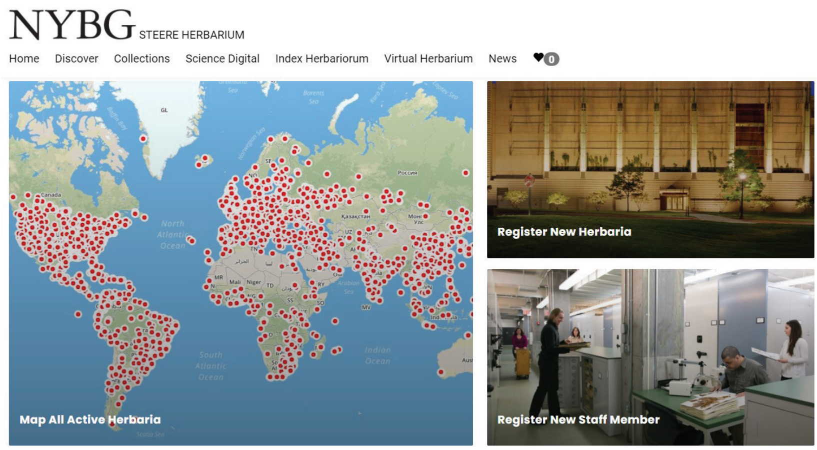
services are a good fit for their purposes. Some ISO standards are specific for biological collections:

1. The ISO 9001:2015 standard provides basic guidance to organizations on how to set up a quality management system with commitment from senior management to support the collection (ISO, 2015).
2. The international biobanking standard ISO 20387:2018 provides additional guidance for a culture collection. The guidance in the biobanking standard solidifies what culture collections have been working toward based on best practices of culture collections around the world (ISO, 2018).
3. A standard that is currently under development, ISO/TC 276 Biotechnology, is intended to bring standardization to the field of biotechnology for biological data and sequence information, which will help to support the information that a culture collection is able to provide. These standards will assure that biological data are accurate and appropriately linked to the specimens and that they are disseminated in correct formats and for appropriate uses.

In specific cases, such as when a user needs a stock microbial strain to diagnose a disease, the fit-for-purpose resource may need to be ISO certified. However, in most cases, formalized but non-certified quality control standards and best practices are sufficient assurances of quality. Networks of collections have proven to be particularly effective in raising quality control standards and elevating customer service. For example, the Mutant Mouse Resource and Research Center is a network of four major collections of mutant mice, with centralized ordering and quality control divisions. WFCC lists 23 regional and international networks of culture collections, including the United States Culture Collection Network. Several of these networks, including WFCC, have developed a shared set of best practice guidelines.

The Global Biological Resource Centre Network, an OECD-endorsed pilot project, is one particular project that may be a useful model for enhancing quality control in U.S. living stock collections. Arising from networked European Union (EU) collections, the Global Biological Resource Centre Network led to the creation of the EU-funded Microbial Resource Research Infrastructure (MIRRI) project (2012–present). With funding expected to exceed 1 million euros per year, the initial aims for MIRRI were to advance collections to become biological resource centers (BRCs), network BRCs, and interact with the user and regulatory communities (Stackebrandt et al., 2015). Examples of activities that benefited EU collections include the development and implementation of quality control practices to become ISO9001:2015 certified; cooperation on databases, websites, and marketing; and gauging and enhancing user satisfaction. These activities have helped advance the EU bioeconomy, supported innovation, promoted global cooperation, and helped both governments and collections meet global requirements such as the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from

BOX 4-6
Toward a Universal Collections Registry: The Example of the Index Herbariorum



Snapshot of the William & Lynda Steere Herbarium of the New York Botanical Garden website illustrating the interactive maps, a picture of the herbarium from the outside, and the staff working inside the building. See http://sweetgum.nybg.org/science/ih.

Every good ecologist knows that to preserve a species, you need to know what it is, where it is found, and how it interacts with other species in its environment. The same, as it happens, is true of biological collections.

their Utilization to the Convention on Biological Diversity and biosecurity protocols. Comparable efforts in the United States would likely have a similar cost but have not yet been implemented.

Establishing a National Registry of Biological Collections

A registry of the biological collections held in U.S. institutions would enable users to discover and contact collections with holdings of potential interest, thereby increasing access to them. It would also improve the ability of biological collections with geographic, temporal, or other commonalities to find one another for potential collaboration and to identify the most relevant collections to include in such collaborations. A comprehensive collections registry could also facilitate an assessment of the infrastructure needs of all U.S. collections and perhaps help prioritize grant funding for infrastructure improvement. It might also facilitate the response of the collections community to emergencies caused by

Since 1935 the Index Herbariorum (IH), a directory to the world's herbaria, has been the go-to place for information about the world's herbaria. The IH was begun in the Netherlands, but the New York Botanical Garden assumed responsibility for managing it in the mid-1970s. It became an online resource in 1997.

Keeping track of the world's herbaria is not an easy task—with every week there are new herbaria, new staff, and new holdings to register, and there are closures or mergers of one herbarium with another yearly. But it is essential for botanists and other scientists and researchers to know where to find and how to contact the curators and staff of the roughly 3,300 herbaria in the world today. Collectively, these herbaria contain almost 400 million specimens; the IH also lists approximately 12,000 associated staff, including curators, managers, and other biodiversity experts. Each entry in the IH (http://sweetgum.nybg.org/science/ih) includes the herbarium's physical location, web address, contents, and history as well as the names, ages, contact information, and areas of expertise of associated staff.

The information contained in the IH allows herbarium staff not only to address their shipping boxes correctly, but also to find individuals who can identify or evaluate specimens and to find partners for specimen exchange. Biodiversity scientists use the IH to find previously collected specimens that are pertinent to their studies. Scientific journals require the use of the IH codes in the citation of specimens examined and the designation of type specimens in the description of new species. Collecting permits for national parks and other protected federal lands require that plant and fungal specimens collected on these sites be deposited in IH-listed herbaria. The U.S. Fish & Wildlife Service uses the IH as a resource for determining whether an institution should be granted a permit to house endangered species (a Convention on International Trade in Endangered Species permit). The IH is also used by the Department of Homeland Security to find specialists for the identification of unknown specimens confiscated at U.S. Customs and Border Protection sites.

natural disasters or infrastructure failure or to anticipate the orphaning of collections and advise on the best options for a collection transfer when needed. The ability of those in charge of a biological collection to compare their collection with others would help inform strategic planning. More communication among collections could lead to the development of more community-wide standards in curatorial practice and data management. The herbarium community already has such a curated registry, Index Herbariorum, which could be a model for a registry that includes all collections (see Box 4-6). WFCC also has a global registry of culture collections. However, the WFCC registry is an opt-in system, which leads to underrepresentation of some countries, underscoring the need for clear criteria for including a collection and for the active curation of registrant information.

Several previous attempts to create a global index of all collection types (e.g., GRBio and Biorepositories.org) have failed to produce a comprehensive registry that is regularly updated with current information. iDigBio recently created a static list of collections in the United States, drawing on Index Herbariorum, previous lists, and information gleaned from institutional websites, but this is only a start (see Box 4-7). In collaboration with Index Herbarium, the Global Biodiversity Information Facility (GBIF) plans to create the platform for a comprehensive worldwide biodiversity collections database (Hobern et al., 2019). A registry of U.S. collections could use the GBIF cyberinfrastructure but will still require a significant campaign of outreach to collections institutions to provide data and develop the tradition of updating the index as holdings and staff change. Similarly, an interest group of the Taxonomic Databases Working Group is working on a collections descriptors data standard for the description of collection-level metadata,[19] which will provide a framework for specimen metadata that needs to be collected in order to provide a full assessment of the strengths and opportunities that these collections may provide to research and education.

CONCLUSIONS

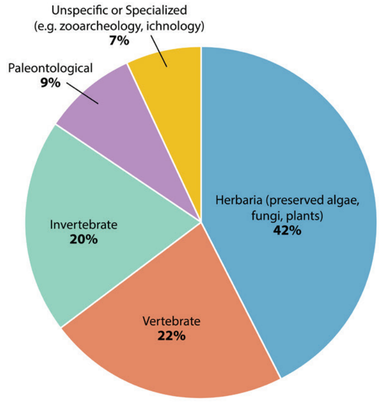
As long as research, education, and the preservation of natural heritage are national and global endeavors, it will be imperative that the infrastructures for biological collections at both the individual and collective levels are improved and maintained. Given the negative consequences of the nation's research efforts if biological collections are limited by poor infrastructure or perhaps are lost altogether, it is crucial that proactive measures be taken to strengthen the physical, digital (see Chapter 5), and intellectual (see Chapter 6) assets that support the long-term quality and curation of specimens and their associated data.

Owing to the diversity of collection types, there is no one-size-fits-all list of physical infrastructure requirements. The assessment of infrastructure needs must take place at the individual level. Unfortunately, most biological collections do not have sufficient resources for preventive maintenance or basic upgrades for existing infrastructure and technologies, let alone for major renovations or new facilities. Thus, biological collections would benefit from individualized strategic plans to outline how day-to-day needs will be met, including issues related to preventive maintenance and quality control, and also how to develop or expand infrastructure to meet future scientific needs.

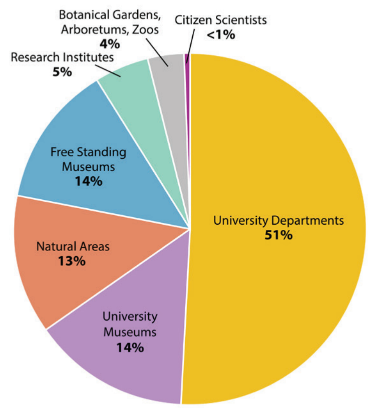
Some aspects of infrastructure will benefit from shared community standards. This is particularly true for quality control for living stock collections, for which consistent genetic identity of the specimens within and between stock collections is crucial for research. Smaller living collections may not be able to afford the staffing and other costs for a quality assurance and quality control system and may not be able to meet the ISO guidelines. Strategic planning will be an important tool in guiding those living stock collections to adopt and maintain nationally accepted quality control standards. In addition, a community approach to developing "next best" protocols and best practices that can be implemented in a way that is commensurate with the available budget could allow such collections to distribute accurately identified pure cultures. For all collection types, a community approach as an additional layer to strategic

[19] See https://www.tdwg.org/community/cd.

BOX 4-7
iDigBio Listing of Biological Collections



Type of Biological Collection



Affiliation

Natural history specimen collections represent a vast distributed network of information on the biodiversity of our planet. Estimates of the total number of specimens held in U.S. collections range from 800 million to 1 billion. The most comprehensive listing of collections for the United States is the iDigBio Collections catalog (https://www.idigbio.org/portal/collections), which lists approximately 1,600 natural history collections in the United States associated with 729 different institutions. This list includes a large variety of collections of different sizes and affiliations but is not complete and particularly underrepresents small, regional collections and private, personal collections. The charts above show a breakdown of the biological collections in iDigBio by type (generalized categories of taxonomy) and affiliation. "University department" refers to collections held in laboratories or other spaces allocated to an academic department of science.

planning could create mechanisms to pool resources and facilitate the development of best practices and training and ensure that the leadership of biological collections is well equipped to implement those collections' plans.

Biological collections infrastructure also needs to grow in order to keep up with the advance and evolution of scientific research itself. The urgent need to continue collecting will require NSF and other funding institutions, as well as institutions whose mandates include collecting or generating new types of research specimens, to acknowledge and address growth as an important and necessary component of biological collections in the 21st century. Such institutions will also need to acknowledge the ongoing demands that collections growth places on infrastructure—demands that can only be ameliorated through infrastructure support and improvements. Such an acknowledgment will require the development of clear guidelines and metrics for growth.

It will also be important for the infrastructure needs of individual biological collections to be integrated into larger infrastructure initiatives of their host institutions and the community as a whole, especially those initiatives aimed at developing state-of-the-art research hubs that meet the needs of a dynamic scientific enterprise. In such endeavors, the institutional staff charged with maintaining institutional research infrastructure will need to understand the particular needs of biological collections in terms of environmental controls and other sensitivities that can affect preventive maintenance and quality control.

Finally, consideration needs to be given to biological collections as a shared and distributed scientific resource for the nation. This will require a consortium to create community-wide mechanisms to pool and share resources. Establishing a registry of biological collections in the United States will be an important step toward cultivating national attention and perspective. Such a registry could be used to conduct periodic community-wide assessments of infrastructure needs. NSF has the opportunity to provide the backbone to cultivate partnerships so that collections across the spectrum are involved in contributing to an emerging consortium.

RECOMMENDATIONS FOR THE NEXT STEPS

Recommendation 4-1: The leadership (directors, curators, and managers) of biological collections should assess and define the infrastructure needs of their individual facilities and develop comprehensive strategic plans in accordance with those needs and their strategic missions. The strategic plans should outline approaches to:

- continually address ongoing preventive maintenance and, in the case of living collections, quality control requirements; and
- improve and potentially build new infrastructure, both of which actions are particularly important if collections growth is a component of the strategic mission.

The strategic plan should be revisited every 3 to 5 years to ensure that it continues to meet the evolving needs of collections and their users.

Recommendation 4-2: Biological collections should take advantage of existing training opportunities and collaborative platforms at the national and international levels, such as those offered through the International Society for Biological and Environmental Repositories and the Organisation for Economic Co-operation and Development certification programs, especially as new aspects of the work evolve, such as regulations compliance, data management, and new techniques and materials for collections storage and documentation.

Recommendation 4-3: Professional societies, associations, and coordination networks should collaborate and combine efforts aimed at addressing community-level infrastructure needs of the nation's biological collections, including:

- develop a platform to pool and share resources such as strategic plans, best practices, and training opportunities so that these can serve as resources for the broader biological collections community;
- develop and implement strategies to adopt quality control programs to improve uniformity among living stock collections and ensure the availability of high-quality biological resources that best fit the needs of the user;
- create a national biological collections registry to document the location, size, and holdings of the collections in the United States. The registry should be curated and updatable. In addition, proactive processes to identify collections should be established, ensuring that collections of all types are well represented in the registry; and
- use the national registry to conduct periodic community-wide assessments of needs to inform the development of both individual and community-level strategies to maintain and upgrade infrastructure.

Recommendation 4-4: The National Science Foundation (NSF) Directorate for Biological Sciences should continue to provide funding support for biological collections infrastructure and expand endeavors to coordinate support within and beyond the Directorate. Specifically, NSF should:

- support new and improved infrastructure to accommodate the pressing needs created by continued collections growth;
- require a specimen management plan for all research proposals that includes collecting or generating specimens that describes how the specimens and associated data will be accessioned into and permanently maintained in an established biological collection; and

- facilitate the creation and support of an independent consortium to develop collaborative platforms and mechanisms to pool and share resources for strategic planning, preventive maintenance, quality control and assurance, collections growth, establishing a national collections registry, and other community-level assets.

5

Generating, Integrating, and Accessing Digital Data

Throughout most of their history, biological collections and the physical specimens they contained were explicitly linked to the physical locations where they were housed. These biological collections consisted of specimens and their accompanying data in written records, and to access the collections users had to travel to the collection or receive specimens through the mail. That is changing now, however, as increasing numbers of biological collections have been digitized. This digitization[1] of specimen data, combined with the cyberinfrastructure[2] that underlies how digital data are stored, managed, and used, has fundamentally transformed the biological collections community (Ball-Damerow et al., 2019; Hedrick et al., 2020) and the work of researchers who rely on biological collections as digitization makes possible the remote examination of biological collections and greatly enhances their discoverability and usefulness.

A key component of digitization has been the development of collection databases that provide digital specimen data to aggregated data repositories, producing a global biodiversity infrastructure. Online data repositories democratize access to digital specimen data, making possible new avenues of scientific inquiry, promoting the multiplication and expansion of research collaborations and community networks, and providing a greater range of educational and training opportunities (Lacey et al., 2017; Monfils et al., 2017). A robust cyberinfrastructure can also facilitate evaluation and the development of metrics for assessing the diversity of biological collections and their impact on research and education (Meehan et al., 2019) (see Chapters 2 and 3).

Biological collections have driven increasingly integrative and collaborative science—with the potential to address a wide variety of problems from disease, such as coronavirus disease 2019 (Cook et al., 2020), to species responses to climate change (Meineke et al., 2018a)—which in turn has intensified the need for greater access to high-quality digital data. Over the past decade, a wide range of advances in the process of generating digital data of all kinds and building the cyberinfrastructure for biological collections has emerged. However, the robust cyberinfrastructure that the biological collections community requires has yet to be fully realized. This chapter focuses on the challenges of and strategies for

[1] The conversion of textual, image, or sound-based specimen information to digital formats.

[2] Cyberinfrastructure, a term first used by the National Science Foundation, encompasses the computing systems, repositories, advanced instruments, software, high-performance networks, and people that enable and support data acquisition, storage, management, integration, mining, analysis, visualization, and distribution (adapted from Stewart et al., 2019). See https://scholarworks.iu.edu/dspace/handle/2022/12967.

advancing the accessibility and integration of digital biological collections for research and education.

CURRENT STATE OF DIGITIZATION, DATA, AND CYBERINFRASTRUCTURE

Digitization: An Evolving Process

Biological collections encompass a diverse array of specimen data that span biological, physiological, temporal, and spatial features of the specimens. Digitization is the process of converting these analog or printed specimen data from specimen labels, field notes, card catalogs, ledgers, genetic sequences, images, audio, and video recordings, and more into digital representations. Digitization helps preserve the long-term integrity of specimens by allowing researchers to inspect metadata and digital images without having to access and physically handle the specimens while opening new avenues of data-driven research (e.g., ecological niche modeling). The biological collections community has spent decades digitizing specimen data to increase their visibility and accessibility to researchers, educators, and the general public. In fact, the digitization of specimens and associated materials and the uploading of these digital data into online platforms has long been a requirement for funding programs such as the National Science Foundation (NSF) Living Stock Collections for Biological Research program and its successor, the Collections in Support of Biological Research program, among others. In 2010 the National Evolutionary Synthesis Center workshop outlined a vision and strategic plan for a Network Integrated Biocollections Alliance to "document the nation's biodiversity resources and create a dynamic electronic resource that will serve the country's needs in answering critical questions" (NESCent, 2010, p. 2). At that point, it was estimated that only approximately 10 percent of all specimens in natural history collections worldwide had been digitized (Page et al., 2015). NSF responded to elements of the NIBA plan by establishing the Advancing Digitization of Biodiversity Collections (ADBC)[3] program, which funds digitization efforts that coalesce around scientific questions or themes through extensive collaborative networks, called thematic collections networks (TCNs), overseen by the national coordinating center for these efforts, Integrated Digitized Biocollections (iDigBio).[4] iDigBio now hosts more than 121 million digital specimen records, the majority of which were largely unavailable to users 10 years ago. Based on iDigBio's digitized holdings compared with estimates of specimens held in U.S. collections, it is now estimated that about 30 percent of all natural history specimens in the United States have been digitized. However, there is still a long way to go until all collections have been digitized, particularly given the challenges posed by certain types of collections and the need for a workforce with both curatorial and data management skills. However, thanks to recent efforts, research using

[3] See https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503559.
[4] See https://www.idigbio.org.

natural history collections data, as measured by citations in publications, has increased dramatically over the past decade, reflecting the increasing number of digitized collections (e.g., Ball-Damerow et al., 2019; Heberling et al., 2019) (see Figure 5-1).

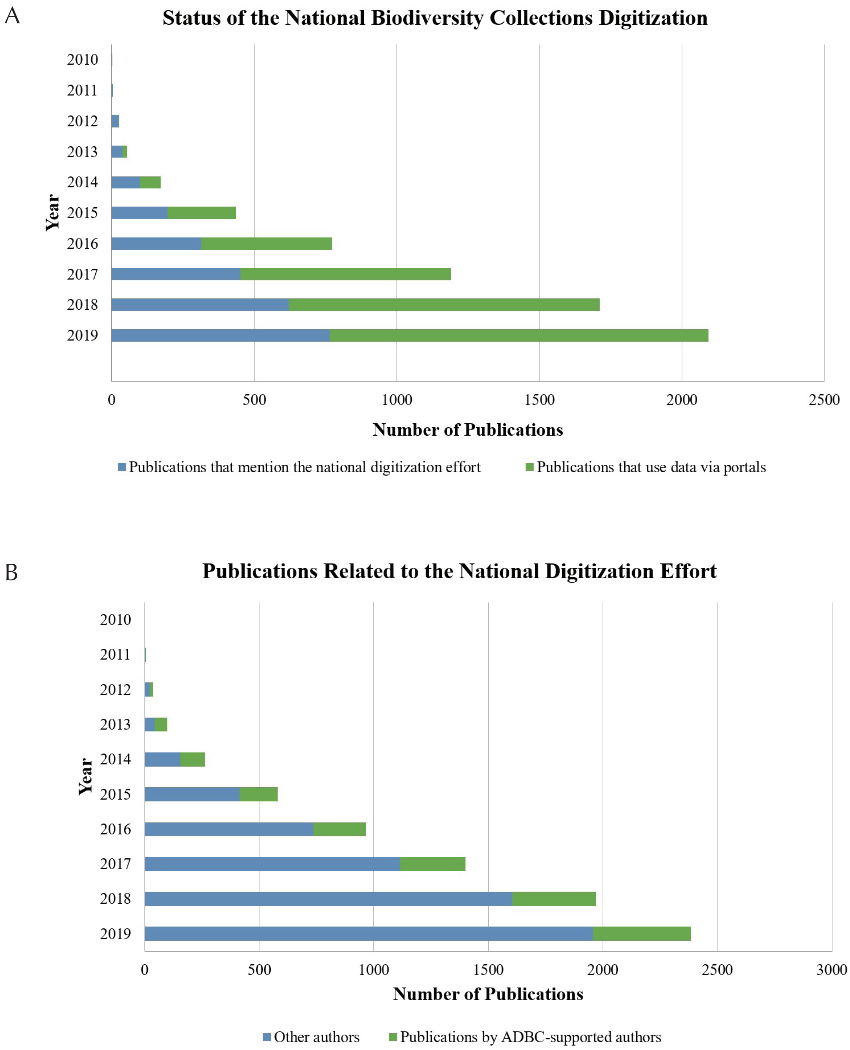


FIGURE 5-1 Publications using digitized natural history data provided and/or served by the National Science Foundation–supported Advancing Digitization of Biodiversity Collections (ADBC) program, 2010–2019. (A) Cumulative number of publications that reference the national digitization effort versus those that use data served by iDigBio and related portals. (B) Cumulative number of publications authored by ADBC-supported investigators versus those authored by the larger community.

The development of digitization workflows (e.g., Haston et al., 2012; Karim et al., 2016; Nelson et al., 2012, 2015; Tulig et al., 2012) over the past decade, coupled with an emerging community of practice among collections professionals, provides a roadmap for accelerating the pace of digitization in the coming decade if sufficient funding can be made available. These digitization workflows provide institutions that house biological collections with guiding principles that can be adapted to their varied needs, collection sizes, and capabilities. Additionally, workshops organized and sponsored by iDigBio[5] and others have made digitization more widely adopted, better understood, and more efficient across the natural history collections community. Living and natural history collections follow the same general digitization workflow (see Figure 5-2), with all collections providing data on the source of the specimen, date of sampling, the collector, and other attributes of provenance. However, the workflows will differ between collections due to the unique digitization priorities of each collection and the varying needs of their respective research and end-user communities.

Rapid technological advances in digitization and cyberinfrastructure have allowed a large amount of historical data to be converted into digital representations over the past 20 years. The current digitization process for existing specimens typically involves hand-entering primary data from a specimen label, field notes, card catalog, or ledger into a database, which can be time-consuming. As described later in this chapter, numerous attempts have been made to speed up this process while preserving the quality of the digital data produced. The pace of digitizing newly acquired specimens, on the other hand, is much more rapid. Specimen data are increasingly "born" digital—directly produced in digital format (e.g., GPS locations, digital spreadsheets, nucleic acid sequencing, three-dimensional images, computer tomography, etc.), which drastically reduces the amount of time required to create specimen records and integrate and share them online.

Toward Accessible and Integrated Data

Digital data from biological collections can be organized into one or more datasets that are collectively stored in local databases. At the local level, digitized collections can be easier to manage than non-digitized collections and may improve the ability of the collection managers to provide access, respond to requests, physically manage space, and allocate budget resources. Digital collection databases can be published and then accessed through online thematic, taxonomic, or geographic data portals, aggregators, and catalogs. Often, the biological collections community uses portals and aggregators interchangeably. In this report a *portal* is defined as the online platform that allows users to perform advanced searches on the published collections found therein. This could be a local portal to an individual collection or a portal of aggregated collections. An *aggregator* is the cyberinfrastructure that gathers and compiles data from published collections and makes them searchable through portals. A catalog is

[5] See https://www.idigbio.org/content/workflow-modules-and-task-lists.

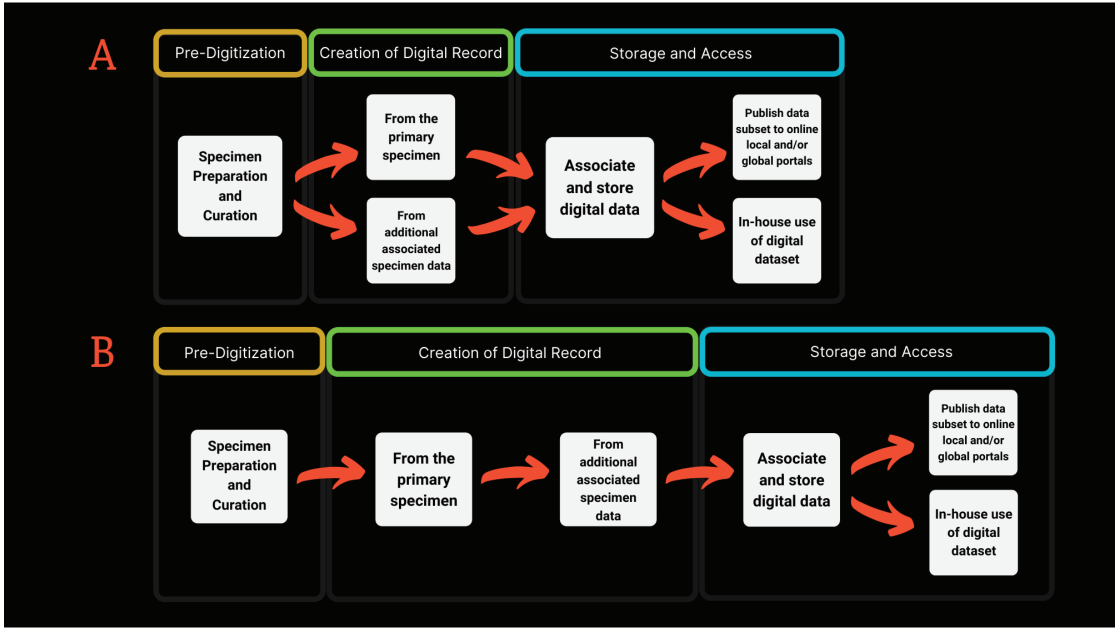


FIGURE 5-2 Generalized digitization workflow. Common pathways for digitizing specimens in living and natural history collections are shown. The workflow begins with curation and preparation of specimens. Thereafter, digitization starts with primary and additional associated specimen data, shown above occurring in two different pathways: (A) the creation of the primary digital specimen record and digitization of the associated data occur concurrently; and (B) the creation of the primary digital specimen record happens first (e.g., scanning of herbarium sheet to digitize record from label), followed by digitization of the associated data, sometimes at a much later date. Specimen data are then associated, typically in a relational database, and stored, ideally, on a server (in-house or cloud-based). Once specimens and their associated data have been digitized, the digital datasets can then be used in-house (e.g., tracking loans and users) and, increasingly, more globally by inclusion in external aggregation sites (e.g., the Global Biodiversity Information Facility, the Global Catalogue of Microorganisms), where they can be discovered and used by a wide range of user groups (e.g., ecologists, policy makers, educators).

similar to a portal but is a term mostly used by the living collection community. Catalogs enable users to search, request, or buy specimens and materials, facilitate the collection of fees, and provide information on shipping permits, compliance with regulations, and user registration unique to living collections.

A major global portal for natural history collections is hosted by the Global Biodiversity Information Facility (GBIF), while iDigBio hosts a portal for collections primarily based in the United States, and the Atlas of Living Australia (ALA) and the Distributed System of Scientific Collections provide portals to Australian and European collections, respectively. There are also project-based portals (e.g., TCN portals, such as the SouthEast Regional Network of Expertise and Collections) and taxonomic portals (e.g., Vertnet, Fishnet, EntoWeb, iDigPaleo). The data that are available via major portals are based on common standards (e.g., Darwin Core Standards[6] or the Access to Biological Collections Data[7] schema). These standards help data providers share specimen data using a common terminology of fields, controlled vocabularies, and data classes that describe the taxonomic identity, collecting event, locality, collectors, geological context, and specimen attributes as well as various kinds of media (Wieczorek et al., 2012). The use of these common standards facilitates the computerized aggregation of data from multiple types of collections and the integration of specimen data with other sources of information. Users are able to search data, download results for further analysis, and integrate the downloaded data with other resources, such as environmental data (e.g., temperature, precipitation, etc.). New standards to allow for the incorporation of additional properties, called extensions, are continually being developed by the global community.

Living stock collections serve as specimen repositories and data providers for members of the research community, who interface with these collections through online databases, catalogs, and aggregators. One such centralized aggregator, the Global Catalogue of Microorganisms (GCM),[8] hosted by the World Federation for Culture Collections and managed by the Chinese Academy of Sciences, facilitates the access and sharing of microbial living stock collections along with their associated data. Online platforms provide information such as available strains, genes and alleles, and genome sequences with functional annotation for the acquisition of research material. Standardized abbreviations for genes, alleles, and depositors and coordinated genome sequencing and annotation projects help make these data useful to the user community (Jarret and McCluskey, 2019). In the GCM, users can locate desired strains along with the associated metadata (e.g., date of isolation, geographic origin, growth conditions, and medium, etc.). Users can add strains to a shopping cart if they wish to acquire them for research, and by putting a strain in the cart a user is linked directly to the source collection, from which the specimen(s) can be requested. Many databases of individual microbial collections are interoperable due to the efforts of projects such as the Common Access to Biological Resources and

[6] See https://dwc.tdwg.org.
[7] See https://www.tdwg.org/standards/abcd.
[8] See http://gcm.wfcc.info.

Information consortium[9] or the now-defunct StrainInfo (Verslyppe et al., 2014), which helped build common datasets based on specific data standards and formats.

Cyberinfrastructure in Support of Biological Collections

Biological collections may offer solutions to various major societal challenges relating to biology and the environment, from the emergence of new pathogens or the need for new antibiotics to the response of species to climate change, but this is possible only if the data can be accessed, aggregated, and analyzed effectively (Cook et al., 2020; Fontaine et al., 2012; Rocha et al., 2014). Following FAIR data principles (i.e., data that are findable, accessible, interoperable, and reusable [Wilkinson et al., 2016])—and the TRUST principles for digital repositories (i.e., repositories that promote the principles of transparency, responsibility, user focus, sustainability, and technology [Lin et al., 2020]) will require a robust cyberinfrastructure. As the digitization of biological collections continues to create large and diverse datasets, an effective cyberinfrastructure will need to incorporate mechanisms to improve access to an ecosystem of digital repositories and enable the integration of diverse types of data. Recognizing the need for a more robust cyberinfrastructure, the Earth science community established EarthCube in 2011 with NSF funding from both the Directorate for Geosciences and the Office of Advanced Cyberinfrastructure of the Computer and Information Science and Engineering Directorate at NSF.[10] Collaborative projects with the biological collections community (such as enhancing Paleontological and Neontological Data Discovery API[11] and Earth-Life Consortium[12]) as well as products resulting from EarthCube have been recommended for adoption by the biological collections community (e.g., Hobern et al., 2019). A similarly broad, community-level endeavor has not yet taken place between the biological science and computer science communities, but the timing is right, given the past decade of focused digitization.

For any local digitization effort to be successful, individual collection-holding institutions need a basic desktop computer and access to server infrastructure in order to house collection management system (CMS) databases, image repositories, and the necessary software for data publishing. Collections also require a workforce skilled in data management as well as collections curation and taxonomy. Both the natural history and living collections communities are using a large number of unique CMS databases that range from simple spreadsheets to more sophisticated systems that allow for database management and data manipulation, such as feature-rich SQL or Oracle-based systems (Arctos, Collections Space, Specify, BRAHMS, Axiell EMu, BioloMICS, GRIN, etc.) with extensive data models, collection management, and publishing capabilities.

[9] See http://www.cabri.org.
[10] See https://www.earthcube.org/info/about.
[11] See https://www.earthcube.org/group/epandda.
[12] See https://www.earthcube.org/group/earth-life-consortium-elc.

Data publishing increases the discovery of specimens for traditional research uses, for research that makes use of the digital data themselves (e.g., predictive modeling, recording of traits through optical character recognition of textual notes, or by machine learning from images), for formal and informal education, and for other novel downstream uses. While many institutions do not have the resources in house to install and maintain the necessary cyberinfrastructure to run a collections database and make their data available online, hosting services provided by web-based collection management packages and community-based solutions provide the cyberinfrastructure and technical expertise necessary to facilitate the digitization and publishing of these collections.

CHALLENGES

Realizing the promise of the digitization revolution will require overcoming a number of challenges. On one hand, there is an extensive community-wide backlog of specimens and associated materials that need to be digitized, creating gaps in our knowledge about the world's biodiversity and missed collaboration opportunities between researchers. On the other hand, the multiplication of shared databases that vary in data quality and format and the proliferation of data aggregators and repositories can lead to an unnecessary duplication of effort, data disintegration, and limited data usability. Mass digitization is exposing digital data to an ever-increasing diversity of users for a myriad of uses, resulting in an increasingly complex digital landscape. Addressing these challenges will require the development, support, and maintenance of robust and coordinated cyberinfrastructure that provides for the ever-increasing needs of the world's biological collections.

Dark Data

While the majority of data generated today are immediately digitally captured, historical collections typically have a backlog of data that have yet to be digitized. The digital revolution and the increase in the accessibility of digitized specimen data have been so profound that undigitized collections are now referred to as "dark data"—referring to the fact that they are essentially unavailable for modern scientific study without physical access to the specimens within institutions (Heidorn, 2008). The absence of these specimens from the global and national collections digital infrastructure represents lost opportunities for research and education as well as limits to returns on the investments made by the funding agencies that supported the acquisition of the specimens, even if the research projects that generated the undigitized collections were otherwise successful.

Discipline-Specific Limitations and Biases

Although digitization efforts to date have been transformational for both biological collections and research communities, most U.S. specimens, especially those from taxonomically diverse groups, remain undigitized and unavailable for inclusion in cutting-edge research. The process of digitization can be particularly challenging for some disciplines where specimen labels are obscured or scarce, where taxonomic diversity is high and poorly known, where the type of preservation precludes automated capture of information (wet specimens in alcohol, for instance), or where the availability of historical paper records (card catalogs, ledgers, field notes, etc.) is limited. For example, for natural history collections, it is estimated that more than 50 percent of vertebrate collections (Krishtalka et al., 2016) and 20 percent of herbarium specimens (Barbara Thiers, Director of the William and Lynda Steere Herbarium at the New York Botanical Garden, personal communication, 2020) are digitized and available online, while only 4 percent of entomology collections have been digitized (Cobb et al., 2019), and most invertebrate biodiversity remains unknown or ignored (Di Marco et al., 2017). Plaguing biodiversity research, taxonomic bias[13] also leads to a disproportional amount of dark data for certain collections and resulting discrepancies in knowledge from organism to organism across a wide range of biological fields (Adam et al., 2017; Clark et al., 2002). Multiple logistical and technical factors contribute to this bias, such as those mentioned above, but regulatory bottlenecks and restrictions play a role as well. Large-scale digitization efforts reveal the extent of century-long sampling and taxonomic limitations and biases and provide insights on how to account for such issues to inform future collecting (Daru et al., 2018; Troudet et al., 2017) and digitization efforts. For some biological collections, certain data fields need to be redacted or restricted and kept dark to protect sensitive information or specimens. This might include the exact geographic location of an endangered orchid or a fossil site on federal land, information and access to particularly virulent strains of biothreat pathogens, and personal identifiers in the case of organisms or samples originating from human specimens.

Project-Based Collections

A potentially large body of dark data lies in project-based collections—a group of specimens or samples collected with a particular purpose (e.g., for a specific research program, project, or survey of a group of organisms in a particular region) but never transferred to a permanent physical repository (e.g., museum collection or biological research center). While these valuable collections could make important contributions to science and society, the key problem is that they typically reside in an investigator's lab, freezer, or office, making them difficult to identify and locate (for more, see Chapter 4). Typically, these collections are not accessioned, digitized, and made accessible to the wider scientific community through national data portals or catalogs. The barriers preventing accessioning into repositories and the subsequent digitization can be diverse. While some projects produce scientific publications that describe their findings and the materials accumulated, researchers may not be willing to share—or may be

[13] The fact that some taxa are more investigated than others.

reluctant to relinquish control of—the specimens in their project-based research and thus be hesitant to contribute them to a publicly available repository or data portal. Even when researchers are willing to contribute their specimens and data, sometimes collections simply do not have the capacity or the resources to entertain such requests because of limited space and inadequate funds for accessioning and digitizing the specimens. Some project-based collections may not be suitable for incorporation into a permanent collection or digitization because of the recipient institution's acquisition policies and guidelines (e.g., strategic growth, accessioning limitations, permits) or an inability to assess the value of a project-based collection and its benefit to the institution.

Private collections are also difficult to find. While outside the purview of this report, these private collections may hold essential data for documenting biodiversity, which may eventually be accessioned in public collections. Although the number and holdings of private collections in the United States are unknown, a recent survey in Europe found that private collections there may make up as many as 33 million specimens (Willemse et al., 2019). There are obvious issues concerning data quality and the willingness of these private collection holders to digitize and publish the data associated with the collections, but this information from Europe suggests that U.S. private collections may be a particularly valuable source of biodiversity data currently invisible to the research and education communities.

An Inefficient Data Pipeline

Currently, each online portal or aggregator collects a copy of a collection's data published on a local database and ingests, normalizes, aggregates, and re-publishes this copy online. However, the current data publishing landscape lacks a streamlined and standardized pathway for carrying out these steps. For instance, if a collection shares its data with multiple aggregators, each aggregator may serve slightly different versions of the same record because they each have different publication schedules and different displayed fields for the specimen data. This publishing process and subsequent data verification steps (taxonomic and geographic verification, data cleanup, annotation, etc.) result in a massive duplication of effort by the aggregators and they each reconcile the specimen digital data while also creating confusion on the part of data users presented with multiple, yet slightly different, copies of the same data. Thus, while large amounts of data are appearing in portals, effective access to these data requires informatics expertise to remove duplicates prior to research use. As a consequence, some researchers and educators who may lack sufficient data management skills will rely solely on a single portal rather than exploring other portals for additional data—a practice that likely limits the number and possibly the diversity of the specimens obtained from a search. Furthermore, there is no effective mechanism in the current data publishing model for effectively and efficiently returning user annotations of data to the original data providers for incorporation into the data stream, resulting in a complete loss of this effort on the part of users of the data for the collections community. Leading aggregators such as GBIF, iDigBio, GCM, ALA, and others recognize the problems of duplicate records and version control (Hobern et al., 2019) as well as the inadequate methods for annotation, but so far they have been unable to develop either a short-term fix or long-term solutions.

Variability in Data Quality and Format

As the quantity of digital data dramatically increases, the presence of incomplete data, data of questionable quality, and a lack of standardization limit both the roles that biological collections data can play in research and education and their usefulness. Issues such as incomplete data records and inaccurate or poorly transcribed data are ubiquitous and lead to limitations on the use of specimen digital data. For instance, an investigator searching on higher-level taxonomy, such as plant family, would miss records for which this information has not been recorded at a higher level but only at a lower one. Studies attempting to quantify the timing of animal migration or plant flowering would be severely hampered by a lack of specific temporal information. Some disciplines (e.g., botany) have used skeletal records[14] as an initial step in digitizing specimen records in order to save time (Nelson et al., 2015; Rabeler, 2015), but while this method opens up a large number of records for discovery, some of the information in these records has yet to be completely digitized, meaning that certain fields of information are not readily available for research. Similarly, although some disciplines have made great strides in community georeferencing[15] endeavors, such as the NSF-funded Mammal Networked Information System, ORNIS (A Community Effort to Build an Integrated, Distributed, Enriched, and Error-checked ORNithological Information System), HerpNet, and Fishnet collaborative projects (Chapman and Wieczorek, 2006), many specimen records are not yet georeferenced and are thus unavailable for spatial analyses such as ecological niche modeling and species distribution analyses (Bloom et al., 2017; Seltmann et al., 2018). Other locality records may never be able to be georeferenced because of historical limitations in the precisions of their locality information.

Data transcription errors and a suite of taxonomic naming issues (Nekola et al., 2019) create a variety of other issues. For example, the rate of errors in geospatial designation or taxonomic classification, either through synonymy or misidentification, has been estimated to range anywhere from 5 percent to 60 percent (e.g., Goodwin et al., 2015; Nekola et al., 2019). Without adequate taxonomic resolution, taxonomic incongruencies can result in incomplete species distribution and trait information. In addition, a lack of adherence to standardized terminology and controlled vocabularies, as well as limitations of or incorrect mappings to Darwin Core fields, has led to various problems in data analysis. For example, attempts to compile information on all "females" of a species are hampered by the numerous variants of this term in the sex field—F, Female, female, etc. (e.g., 2,800 distinct values appear in the sex field in VertNet; see Guralnick et al., 2016). Approaching the issue at the source by standardizing

[14] A basic set of data per specimen (Nelson et al., 2015).
[15] The process of converting a text-based description into a geospatial coordinate.

and controlling vocabulary in local collections databases and providing common names for organisms would increase usability, but a consensus on taxonomy, terminology, and common names among scientists, which will be needed in order to enable such functions, is still elusive in some disciplines.

Limitations Affecting Data Usability

Once published to a portal, digital datasets require collections professionals to curate and maintain their quality, just as physical specimens require specialized care. Inadequate maintenance of these datasets can severely impair the use, value, and impact of biological collections data in research and education. Both local and community-level mechanisms could improve the quality of their data. One challenge is the lack of expertise by collections professionals in evaluating data quality across broad taxonomic distances and types of data, although standardized vocabularies could provide the necessary tools to assess data completeness, quality, and consistency and to increase the fitness-for-use of biodiversity data (Ball-Damerow et al., 2019). Data transcription errors also require correction by individual collections or potentially through community efforts (see Nekola et al., 2019, for a summary). However, digital datasets are often not maintained and updated for a variety of reasons, ranging from insufficient resources and staff turnover to disputes related to intellectual property rights and to a simple lack of understanding that digital datasets are not static, one-off products.

Another factor affecting data usability is the fact that data portals have been developed for different uses and different communities and their interfaces are not always user-friendly for either the public or the research community. Their design has often been an afterthought because the interfaces for most portals are designed with a single purpose in mind and anticipate only one type of user—the research or collections specialist, and not the general public or student users (Hendy and MacFadden, 2014). Thus, although millions of specimen records are available online, the level of technical expertise necessary for accessing them may be too high for some users. Portals that were designed to serve a wide array of data (e.g., GCM and GBIF) also suffer from limited search capability. Fields that are unique to particular collection types (e.g., mutant allele for genetic stock centers or geological data for paleontological specimens) are not searchable, making those data more difficult to discover. Currently, many data portals are available only in a single or a few languages, providing yet another barrier to accessibility and contribution.

Inadequate Methods for Data Integration and Attribution

Realizing the vision of successfully integrating and tracking data from various sources carries many challenges, most significant of which are issues of scale and interoperability. Data integration relies on the unambiguous identification of individual data elements, packets of data, and people through the use of globally unique identifiers (GUIDs), digital object identifiers (DOIs), and open researcher and contributor IDs (ORCIDs) (Page, 2008) as well as the implementation of standardized application programming interfaces and exchange formats (Konig et al., 2019). Despite several attempts (e.g., Güntsch et al., 2017; Guralnick et al., 2014; Nelson et al., 2018), the biological collections community has been unable to agree on a single identifier to describe data elements, though many candidates have been proposed (GUIDs, life science identifiers, uniform resource identifiers, DOIs, Darwin core triplets[16]). Although most collections now use some form of identifier as listed above, there is no centralized system of registration to ensure the uniqueness—and therefore traceability—of these identifiers, and attempts to link data informatically have been only marginally successful (e.g., Guralnick et al., 2014). Because living stock and natural history collections databases were established in parallel using different types of identifiers, integrating them has proved to be quite complex, and these difficulties may preclude opportunities to integrate the data from these resources. The challenge is exacerbated by the differing types of published data not being comparable, by differing expertise, and by the different user communities being served. As a result, tracking the use of biological collections data in research and education still remains largely a manual and time-consuming endeavor. Issues of tracking multiple identifiers and integrating specimen data across databases and portals are exacerbated by the fact that identifiers do not reliably persist through to the products of research created from the use of these specimens (Arbeláez-Cortés et al., 2017; Rouhan et al., 2017). In fact, even the way that specimens are cited in published work is inconsistent, if they are cited at all. This results in a lack of recognition and attribution of the contribution of biological collections to research and education.

Despite all of the challenges described above, electronic citation and tracking of digital specimen records, each with a unique identifier, can provide attribution to local collections and can enable assessment of short- and long-term impact, both locally and nationally. Although digitization of biological collections has provided access to massive numbers of specimen records, the assessment of the impact of this resource has barely begun (Hobern et al., 2019; Lendemer et al., 2020). Few biological collections have the resources or community-based guidance to take the next step in determining the contributions of their collections to the published scientific body of knowledge. For example, due to incomplete or non-unique metadata in GenBank, even the apparently simple task of automatically connecting genetic data from GenBank to voucher specimens in iDigBio cannot be accurately accomplished, although this connection may be established manually for a given collection, as demonstrated more than a decade ago (Strasser, 2008). While technology may offer some solutions, the development of such citation and attribution systems is in the early stages of implementation—see occCite[17] and GBIF citation metrics and guidelines[18] as promising examples—and it will require substantial investment if these are to be implemented on large scales. The problem is compounded by a lack of coordination among the members of the biological collections community and

[16] A concatenation of values for institution code, collection code, and catalog number for a specimen.

[17] occCite is an online tool that enables biological collections to track how their data are being used. See https://hannahlowens.github.io/occCite.

[18] See https://www.gbif.org/citation-guidelines.

by a lack of appropriate resources to develop and implement an assessment of collective impact. Investing in the development of bioinformatics tools and cyberinfrastructure to capture data used in publications and other forms of output could be transformational in making it possible to accumulate national usage statistics and to carry out rigorous evaluations of the impact of both physical and digital resources.

Limited Mechanisms to Support a Cyberinfrastructure That Promotes Collaboration

The diversity of biological collections poses many challenges to the effective development and implementation of a cohesive, adaptable, and sustainable cyberinfrastructure that serves the entire collections community. For example, inherent differences between living and natural history collections such as differing needs and goals, compounded by external factors such as different funding opportunities and requirements, have thwarted collaborative efforts to integrate digital data from these collections. Many natural history institutions with the necessary funding for personnel and technology have been digitizing their collections for four decades (Nelson et al., 2018), but NSF's 10-year, $100 million ADBC program, launched in 2011, has led to even greater strides in digitization and provided access to an ever-increasing quantity of data from natural history collections. In contrast, at present, living stock collections are ineligible for funding through the ADBC program, and, for now, no comparable programs specifically fund the digitization of living biodiversity collections. The immense amount of digital information being produced by current digitization efforts and the data integration challenges outlined above threaten to outstrip the necessary cyberinfrastructure support (storage devices, backup systems, routine maintenance, and technological upgrades). The financial outlay required for these necessary components and additional workforce needs (see Chapter 6) is sometimes not adequately factored into the cost estimates of digitization, so that the infrastructure components and workforce needs are left unfunded (see Chapter 7), with it being necessary to put retroactive measures in place to address the issue in hindsight. Without sufficient investment in these cyberinfrastructure components and support by individual collections, funders, and the community as a whole, the amount of digital data stored, shared, and integrated will continue to be limited for certain collections. However, it is precisely a broadly based, flexible, and robust cyberinfrastructure that could integrate complementary data from living and natural history collections (e.g., microbiome studies, food safety, biotechnology applications, etc.) or other groups of collections.

THE WAY FORWARD

Digitization is increasing the relevance of collections in diverse ways and allowing collections around the world to network their way toward the "global museum" that will seamlessly integrate worldwide collections (Bakker et al., 2020). To date, the digitization of biological collections has proved extremely valuable and successful. The result has been new partnerships for innovative scientific inquiry and learning. Digitization has significantly increased the accessibility and usability of biological collections data for traditional research, for new research of global societal importance, for education (e.g., Cook et al., 2014; Powers et al., 2014), and for an ever increasing and ever more diverse collection of additional applications (for review, see Ball-Damerow et al., 2019; James et al., 2018; Krishtalka et al., 2016; Nelson and Ellis, 2018).

However, if such successes are to continue and multiply, a great deal of work remains to be done. A large percentage of the nation's biological collections have not yet been digitized. Data cleaning exercises, standardization, and the provision of annotation mechanisms will significantly increase the usefulness of both the collections that have already been digitized and those that will be digitized in coming years. Finally, digitized biological collections will be most valuable as components of a highly integrated cyberinfrastructure that provides easy access to the collections, integration among different collections and with data beyond collections (such as environmental data, genetic data, and biodiversity analyses), and a way to enable effective collaboration among the many researchers who work with those collections and among potential users of the data. These steps will make it feasible to fulfill the extraordinary promise of digitized biological collections.

Innovative Approaches to Reducing Dark Data

Given the foundational role that digitization plays in the development of an accessible, useable, and networked scientific infrastructure, it is important that biological collections continue to digitize and to provide data that are of high quality, in a standardized format, fit for use, and broadly accessible. Digitization workflows are currently in place in many communities and institutions, and systematic digitization is set to become more efficient than in the past thanks to ongoing training support by iDigBio and others. The quantity of digital data available for end use is determined not just by the pace at which historical data can be digitized, but also by the efficiency of adding new field-collected materials or project-based collections to permanent repositories and online portals. To avoid contributing to the backlog of undigitized material, the large amount of data associated with these new specimens needs to be "born digital." Streamlining their integration into collection databases and online data aggregators will require collaboration among field collectors, collections professionals, and the informatics community. By building on recent achievements of the collections community, future efforts to digitize most U.S. collections seem feasible, given sufficient time and funding.

Massive digitization efforts to capture and place online not only the metadata associated with biological specimens but also high-resolution images of the specimens themselves, along with videos and vocalizations, have unleashed entirely new areas of study. Thanks to new imaging techniques and technologies, the use of rare or fragile natural history collections is less invasive, and it is possible to carry out detailed examinations of specimen attributes without extensive

BOX 5-1
Eggs Benedictine: Crackless Analysis of Eggshell Composition

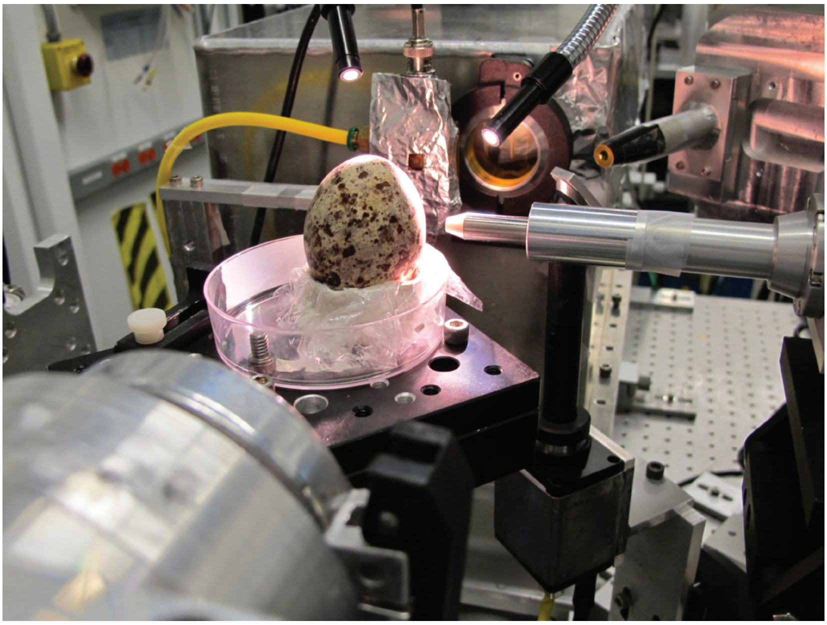


Image courtesy of Monica Tischler, Benedictine Univerity

As organisms grow, they can incorporate numerous signatures from the environment around them into their bodies—including environmental contaminants. Scientists have long used material in biological collections to study changes in these contaminants and their biological effects over time, such as the thinning of eggshells in birds of prey as dichlorodiphenyltrichloroethane levels in the environment increased.

Usually, though, the techniques used to study contaminants in biological specimens result in the destruction of the specimen itself. Eggs are a good

handling of the specimens themselves (see Box 5-1). Sensitive computed tomography (CT) methods of scanning whole organisms and individual skeletal elements capture anatomical features in unprecedented detail and permit precise three-dimensional replication of specimen morphology. Other technological advances have made the digitization of some collection types less time consuming and more efficient (e.g., trays of insects with multiple labels, fluid-preserved specimens, microscopic organisms). Batch processing or automation and the use of optical character recognition (OCR) have shown some success in optimizing the capture of text from specimen label images. The secondary augmentation example: If you want to find out what birds were exposed to in the 19th or 20th century, and you have eggs collected and preserved from that era (egg collecting, or oology, was a huge Victorian craze), you can crush the eggshells and submit them to chemical analysis. But then you do not have an egg anymore.

Monica Tischler, professor of biology at Benedictine University, solved the problem of destroying egg specimens in order to study them by using eggs from the university's Jurica-Suchy Nature Museum and Argonne National Laboratory's Advanced Photon Source (APS) (Lab Manager, 2015). The APS produces some of the most powerful X-rays available, powerful enough to "see" chemical composition in the eggs without destroying them.

"We have eggs dating back 150 years," Tischler said. "Before binoculars were invented and made bird-watching popular, many people collected bird eggs. Then when migratory bird acts were instituted in the late 19th century and made the practice of collecting eggs unfashionable and illegal, many collections were donated to museums" like the one at Benedictine University.

"When birds lay eggs, they excrete contaminants into the egg, and the contaminants in the eggshell reflect blood concentrates of those contaminants," Tischler said. "These specimens represent a window into the past. The problem is that up until this research, all the techniques used to identify the contaminant in an eggshell were destructive. You take the eggshell, crush it, dissolve it in acid, and examine it. It would be unfathomable to destroy these rare eggs for research."

Researchers identified naturally occurring elements such as calcium, iron, and zinc within eggs, but also elements such as manganese, arsenic, bromine, and lead, which can be considered contaminants. "It's a new technique to gain a window into the past to compare watersheds and compare contaminants over time," Tischler explains.

But you have to have the eggs on hand, in this case, thousands of egg specimens amassed by the late Benedictine professors Frs. Hilary and Edmund Jurica, O.S.B., over a period of decades in the early 20th century and later donated to the museum.

of records through georeferencing[19] can be facilitated through the use of online software such as GEOLocate.[20] Specially designed robotic systems that select and image individual specimens or scan whole drawers of specimens and their data are now a reality. The use of convolutional neural networks, a form of machine learning that has been used for species identification (e.g., Carranza-Rojas et al., 2017) and the capture of trait information from specimen images and text such as whether a specimen is in flower or fruit (e.g., Lorieul et al., 2019), is another

[19] Assigning a latitude and longitude to a collection locality (e.g., GEOLocate, Google Earth).
[20] See http://www.geo-locate.org.

area of innovation that could advance digitization (see Box 5-2) and that is ripe for collaboration with computer scientists.

The natural history collections community has begun to use outside assistance in the digitization process in an effort to reduce the amount of dark data. The impact and contribution of citizen scientists and volunteers to the digitization effort have steadily increased through efforts such as Notes from Nature[21] and the Smithsonian Transcription Center,[22] among others. The annual Worldwide Engagement for Digitizing Biocollections (WeDigBio)[23] global transcription event has also galvanized these digitization efforts by engaging a large and diverse set of individuals from varying backgrounds in the digitization process (Ellwood et al., 2018). Although these citizen science efforts were originally designed to assist with the transcription of specimen label data, field notes, and other text (Hill et al., 2012), citizen scientists are extending their contributions to other forms of data capture, such as scoring herbarium specimens for phenological phase. Despite lingering skepticism about the quality of data produced by citizen scientists, it has been found that, when given appropriate instructions, citizen scientists produce data that are on par with specialists (Brenskelle et al., 2020; Catlin-Groves, 2012), and the power of engaging citizen scientists is shown by the fact that the 4-day WeDigBio event in 2018 resulted in more than 50,000 record transcriptions (Ellwood et al., 2018). However, despite the addition of these efforts to existing collections digitization efforts, most of the nation's collections remain to be digitized.

It is important to note that the physical specimen is the nexus for the digitized data associated with it and that it should not be neglected or discarded. Often, the specimens remain the primary source of verifiable biodiversity data, and the curation of the underlying specimens required for such analyses remains paramount, especially if researchers want to later examine the physical specimens after analyzing data from digitized information such as images or genetic sequences. For example, downstream analyses can include the extraction of DNA for the confirmation of species identifications based on analyses of digitized specimens or a simple inspection of specimens for verification and occurrence that might appear anomalous in terms of locality or habitat. As such, digitization is not a substitute for physical specimens, but rather a necessary complementary activity that exponentially increases the usefulness of and provides wider access to the collections of these physical specimens. In fact, evidence is accumulating that use of physical specimens through loans and visits to collections has actually increased with the recent online accessibility of digital records (Vollmar et al., 2010). For living stock collections, continued digitization allows researchers around the globe to locate and acquire an ever-growing number of existing and newly developed model organisms, with the digital data being more of a finding tool and the physical specimen still remaining vitally important. In some cases, such as in the case of destructive sampling or loss of a specimen, the electronic

---

[21] See https://www.notesfromnature.org.
[22] See https://transcription.si.edu.
[23] See https://wedigbio.org.

BOX 5-2
Leveraging Machine Learning to Augment Digital Data Potential

The increasing availability of digitized collections data—textual, geographic, and images—is enabling the application of novel technologies for innovative research. One such application is machine learning, "the science of getting computers to act without being explicitly programmed." Application of machine learning approaches to digital images of herbarium specimens, which are two-dimensional and generally standard in format, is opening doors to new areas of botanical research in ecology, evolution, and agriculture (Soltis et al., 2020, and a special issue in volume 8 of *Applications in Plant Sciences*, 2020). An early application was the development of powerful tools for identifying plant species with an astonishing level of accuracy (e.g., Carranza-Rojas et al., 2017). Likewise, the coupling of digitized herbarium images with machine learning has the potential to revolutionize capture of changes in plant phenology—budburst, flowering, fruiting—across space and time, providing a rich data resource that augments current observation networks of professionals and citizen scientists to assess phenological changes in a changing climate (e.g., Lorieul et al., 2019; Pearson et al., 2020; Willis et al., 2017). An emerging area is the use of herbarium images for scoring so-called "plant functional traits"—those features tied to key ecosystem functions—across species, space, and time for ecological analysis on local and global scales; the application of machine learning to functional trait extraction from images is just around the corner (Shouman et al., 2020; Soltis et al., 2020). Similar approaches are enabling the extraction of trait data from textual information in specimen records—such as body mass, reproductive status, or habitat information—for comparative analysis. Key to all emerging interdisciplinary research uses of digitized collections data is the linkage of collections to heterogeneous data representing environmental, climate, spatial, phylogenetic, and genomic information.

information stored in a database becomes the only record available; this is the case especially for a growing number of microorganism specimens (see Box 5-3), and thus digitization is essential for future studies that aim to understand their biology and evolution.

Increasing Data Visibility

Although digitization and sharing data with online open access data portals continue to provide more data for research and education, vast amounts of data produced through research and collecting endeavors, such as project-based collections data, are still not publicly available. This is particularly prevalent at institutions that lack permanent collections. Making these data public would increase the visibility of the data as well as promote research reproducibility and reduce redundancy. The primary onus of ensuring that data are captured

BOX 5-3
When Electronic Data Become the Only Data

Diverse studies have revealed the existence of large numbers of viruses, bacteria, archaea, and protists (Cai et al., 2019; Coutinho et al., 2019; Ryan et al., 2019) that are not available from any physical collection. In these cases, the only record available is nucleic acid sequences, electron microscope pictures, or the metadata related to the sample and project where they were detected. This is also true for biological collections where specimens or biological material are consumed during research investigation, and the situation is particularly prevalent for environmental samples, such as soil for microbial analysis, marine or riverine water samples, or other new "collections" not yet explored. Without physical material, some collections of DNA cannot be identified taxonomically and therefore cannot be assigned a scientific name. In GenBank it is common to find large sets of sequences that have as source organisms "uncultured sea-water bacterium," which at the time was the best identification possible. In the future, some of these records can play a key role in the definition of new taxa, and the metadata associated with the records represent an opportunity for increased access to data and metadata for an expanding array of biological research questions. For these collections, while common standards and best practices for long-term preservation and curation need to be developed, the biological collections community has the capacity to manage, curate, and integrate new molecular-only collections. For example, some genome projects are aimed at providing a phylogenomic framework to identify otherwise unidentified sequences and understand gene functions (Nagy et al., 2020). In some cases, the increasing number of sequences with physical material that are being lodged with these aggregators can now be used to compare and confirm identification of these non-preserved sequences. As new research is conducted, digital records will need to be updated as physical specimens are re-determined, more organisms are described, and new taxa defined. Some of these "orphan" records with unnamed species could be assigned to these new organisms, but this effort will require careful curation and continuous scanning for taxonomic updates.

and disseminated falls on funding agencies, reviewers, and publishers. The NSF Directorate for Biological Sciences requires a data management plan as part of all research proposals, but while this is a prerequisite for funding for living stock collections, there is no requirement for digitization, publishing, or ensuring the long-term accessibility of specimens and their data for natural history collections. There is thus an opportunity to develop more stringent requirements for managing and archiving specimens and their data as part of a specimen management plan (see also Chapters 4 and 7). Likewise, there is no uniformity in the requirements for data citation in publications through journals. Publishing entities along with their editorial boards (and with pressure from funding agencies) could enact uniform requirements for data citations in order to promote reproducible science as well as to provide the necessary mechanisms for collection attribution.

Tools to Improve Data Quality

Emerging efforts to provide online tools for improving data quality while also facilitating data integration, usability, and accessibility to a broader range of communities hold significant promise in many areas. Both discipline-specific efforts to address data quality and larger-scale efforts by data aggregators provide such opportunities. The aggregator community has a major role to play, with GBIF, iDigBio, ALA, GenBank, and VertNet having already incorporated data quality tests and assertions[24] into their portals, which, in some cases, automatically correct or augment records to enhance their fitness for use (Bouadjenek et al., 2019; Chapman et al., 2020). Most of the changes made as a result of these tests and assertions improve data quality by identifying georeferencing mismatches, genetic sequences that are inconsistent with the literature, taxonomic or geographic anomalies, duplicates, or issues related to data standards or vocabulary. Currently, there is no uniformity in the identification of the errors or in the implementation of the edits across the various aggregators, but recommendations to improve data quality have been proposed (Chapman et al., 2020; Groom et al., 2019). In addition, there is a need to create standardized and consistent mechanisms for feeding these corrections and data flags, or annotations created by users of the data, back to the data providers in order to inform data correction and augmentation at the source. In some cases, annotations and errors found by the users of the data are provided to the data providers in a format requiring corrections to be made individually, one record at a time, which is simply not feasible for large datasets. In the past few years many web annotation tools for eliminating these hurdles have become available (Suhrbier et al., 2017; Tschöpe et al., 2013). Partnering with computer scientists and software developers could lead to the deployment of mechanisms for routing data quality annotations to the data providers and for those annotations to be easily reviewed and integrated into the source data in batches. Machine learning and other forms of artificial intelligence may provide incremental increases in the annotation of certain collections, primarily through text recognition and OCR technologies using images of labels or card catalogs or ledgers. A systematic and standardized approach to improve data quality will result in optimized user experience. Some portals have started to adopt the use of facets, filters, or auto-complete for searching, rather than completely blank entry fields. Such modifications are also steps in increasing the accessibility of collections data to a wider range of users.

Promoting Integration and Attribution

Many national and international organizations have developed standards for collections data management that inform the integrity and format of digitized

[24] A query that looks for problems in a biological collection dataset.

data. These data associated with specimens usually involve a suite of unique identifiers with taxonomic, locality, temporal, and preparation information as well as various collections management–based fields (catalog number, cataloger, etc.). While the fields of information captured may vary by discipline or collection, the widespread adoption of GUIDs would allow for a much deeper and broader integration of data both within and among collections. Collections with a critical body of digitized data based on or derived from the specimens are now interested in linking their basic collection metadata to information such as gene sequences, isotope values, or morphometric analyses. Such linkage will further improve data integration and create better connections between primary specimen records and extended data. Linking the data in this way creates what has become known as the "extended specimen" (Webster, 2017) (see Figure 1-2). Extending specimen data with these resources greatly increases the value of the digitized collection for downstream uses while promoting integrated science (Lendemer et al., 2020; Thiers et al., 2019). A lack of integrated online resources will restrict access to valuable collections information, limiting the uses of the data in research and the potential scientific discoveries related to those data. For maximum use, digital data require integration and interoperability at multiple levels. At the specimen level, data derived from the diverse preparations of each specimen (e.g., skeletons, tissues, parasites, field notes, publications, etc.) need to be connected in order to create full extended or holistic specimens for multidisciplinary applications. In addition, these data need to be integrated with the new data streams derived from subsequent investigations (e.g., GenBank sequences, IsoBank signatures, images, CT scans, viromes, and various -omic data). At the collection level, creating associations among taxonomically disparate specimens to highlight such relationships as tissue–voucher, host–parasite, pollinator–host plant, predator–prey, commensals, and others are crucial for integrative science. At the ecosystem level, many novel uses of biological collections data, such as evaluating species' responses to global change, require integration with other forms of data, such as genetic, observational, trait, environmental, geographic, ecological, and remote sensing data. Such an integration will not only require the collections to be more robust and complete but will also necessitate the creation of interoperable linkages among databases. Some levels of integration of disparate datasets are currently being achieved on a national and global scale through various aggregators and individual museum data management systems, but more coordination between these aggregators and developers is needed to simplify and standardize the landscape.

A cyberinfrastructure for biological collections could enable data integration while also providing annotation tools and a system for attribution of specimen data used in research, education, policy development, or other activities of this scope. Creating such a cyberinfrastructure will require robust technological cyberinfrastructure tools to link data elements and also social incentives that will engage all actors in the data pipeline from collections, to researchers, aggregators, data authority providers (taxonomy), journal editors, and beyond. The promises of data integration and attribution were addressed in a Biodiversity

Collections Network workshop (BCoN, 2018) in which a possible system of identifiers and linkage mechanisms was identified as a solution to better integration and attribution of digitized biodiversity data. For example, a number of systems that are intended to solve various aspects of the integration process are being developed (e.g., GenBank Linkout[25] and Pensoft ARPHA writing tool[26]), but while there are analogous systems in other domains that one could learn from or co-opt (e.g., research resource identifiers of the Resource Identification Initiative[27]), no comprehensive solution has been forthcoming. The more that such technological solutions are implemented, the less the community will need to rely on social solutions where all producers and users of data need to perform linkages manually. The broadened utility of collections data, through integration with other data sources, will eventually increase the use of collections (both physical and digital) and thereby increase the attribution, tracking, assessment of impact, and subsequent advocacy for these resources. Assigning identifiers to downloaded datasets from aggregators would also promote both attribution of data use to the providing institution and reproducible science. For example, GBIF assigns a DOI for a downloaded dataset, but recent research has shown that neither URLs nor DOIs are stable, even over short timeframes, and suggests instead a method of cryptographic content-based identifiers (Elliott et al., 2020). Continued efforts to develop methods for identifiers of datasets to enable data integration, attribution, and reproducible science are needed. One technological solution that could potentially resolve the data integration and attribution problem and that has recently received attention is blockchain (van Rossum, 2017). Blockchain is used most commonly in cryptocurrency where it provides an incorruptible digital ledger of economic transactions. A blockchain-inspired network has the necessary technological components to provide the identification of the various elements of the network while also tracking all transactions associated with each item. The network could take advantage of the existing identifiers commonly used in the collections community (GUIDs, DOIs, ORCIDs, etc.) to effectively identify occurrence records, data downloads, publications, and agents. Transactions such as a change or augmentation of the record by the collection, an aggregator or a user of the collection, a loan or gift of material by a collection to an end user, the lodging of a DNA sequence with GenBank, or the publication of results depending on the use of physical specimens or data could all be digitally recorded by the blockchain network. Each of these individual transactions would be logged by the system and would be traceable and immutable. Further investigation of a blockchain-based cyber-infrastructure could yield innovations for managing and tracking all activities of biological collections.

[25] See https://www.ncbi.nlm.nih.gov/projects/linkout.
[26] See https://arpha.pensoft.net.
[27] See https://www.force11.org/group/resource-identification-initiative.

Developing a National Cyberinfrastructure

As digitization spreads across scientific disciplines and data sharing becomes more common, the development of a flexible, unified, and sustained national cyberinfrastructure would provide greater opportunities to integrate and support disparate digital datasets such as living stock and natural history collections and would facilitate research and educational opportunities. This shared resource would not only serve the needs of the collections communities but also provide a baseline to all biodiversity knowledge. Partnerships and pooled resources may be the key to the development and implementation of a permanent, effective cyberinfrastructure in support of digitization, annotation, integration, and analysis of the nation's collections. Because small collections may have unique holdings that reflect regional species pools or the expertise of present and past local collectors and researchers, making these collections digitally available will be a first step toward greater advocacy, visibility, use, and inclusion in large-scale studies. However, some small collections do not have the resources to manage their own cyberinfrastructure or establish and maintain a portal to store their data or even publish them online. The cyberinfrastructure needs of these collections are in some cases being addressed at the community level through cloud hosting of collections databases (e.g., Arctos,[28] Specify,[29] or BioAware[30]). These web-based collection management packages offer information technology support, which is often not provided in-house by the institution but is necessary to facilitate digitization and publishing of collections. This model has the additional benefits of making data publishing streamlined and making connections to external data repositories more robust (e.g., GenBank, Morphbank, IsoBank, Morphosource, Ontobrowser, DataOne). Data portals such as iDigBio provide global access to digital data from U.S. collections and are therefore a key feature of cyberinfrastructure, but they in turn rely on additional cyberinfrastructure components, such as hardware for servers and storage, an evolving database schema to accommodate innovations in digitization, and a workforce capable of adapting to a rapidly changing data science landscape (see Chapter 6). This type of infrastructure needs to be maintained at the national level, for use by and the benefit of the biological collections community as a whole (see Chapter 8).

A Robust Cyberinfrastructure to Promote Coordination and Collaboration

Connecting data in order to generate shared resources has additional benefits. For example, researchers are increasingly interested in patterns of spatial, environmental, and genetic variation, particularly when evaluating how species might respond to climate change. Data from living stock and natural history collections, environmental databases, the National Ecological Observatory Network, and GenBank would all contribute to addressing these questions, and

---

[28] See https://arctosdb.org/about.
[29] See https://www.sustain.specifysoftware.org.
[30] See https://www.bio-aware.com.

a cyberinfrastructure to support these linkages would enable important new science while ultimately reducing costs through the elimination of duplicated effort. Moreover, the development and deployment of analytical tools and pipelines through unified resources would democratize biodiversity science by allowing accomplished biological specialists who are not well trained in informatics and computer science to address important basic and applied research. Collaboration with a national cyberinfrastructure for life science research, such as CyVerse[31] (funded by the NSF Directorate for Biological Sciences), the Texas Advanced Computing Center,[32] the Extreme Science and Engineering Discovery Environment,[33] and the Data Observation Network for Earth,[34] could provide resources to support biological collections and lead to an enhanced national network of digital data from collections and other relevant repositories by improving accessibility to and linkages among data from different sources. The EarthCube community (see above) could serve as a model for how such a collaboration might be implemented.

A national cyberinfrastructure for biological collections that will support these collections and facilitate their ever-growing base of end users will require collaboration, especially between the collections community and computer scientists and engineers, but also between collections staff from diverse collections types and communities (e.g., natural history and living stocks). Until recently, interactions between these communities have been limited due to a lack of funding and staff availability (see Chapters 6 and 7). However, the effective development and deployment of cyberinfrastructure for biological collections will require both (1) application of recent advances from computer science and engineering in new contexts and (2) innovation of cyberinfrastructure components to meet the unique needs of biological collections and an ever-widening user community (e.g., Heberling et al., 2019). Successful implementation will require an interdisciplinarity that is only beginning to emerge among computer and data science and all fields of biology represented by biological collections. To date, innovations in the development of the world's largest aggregators of data from natural history collections (e.g., GBIF, iDigBio, ALA) and living collections (e.g., GCM) have resulted from close collaborations among biologists, data scientists, and engineers. Moreover, because some computer and data scientists are embracing the data from these biological collections (Chen et al., 2019; Drew et al., 2017), interesting challenges for machine learning and analytical pipelines are being tackled. A similar, although perhaps less appreciated, facet of the situation is that biological collections provide unique and scientifically interesting challenges that could possibly benefit the computer science community, perhaps with extensions to problems outside of collections. NSF's Big Idea "Harnessing the Data Revolution" is certainly relevant to collections data, particularly as both the volume and heterogeneity of data increase and as researchers and

[31] See https://www.cyverse.org.
[32] See https://www.tacc.utexas.edu/-/tacc-a-holistic-approach-to-making-cyberinfrastructure-accessible.
[33] See https://www.xsede.org.
[34] See https://www.dataone.org/working_groups/cyberinfrastructure.

educators are increasingly interested in connecting collections data with other data resources, from environmental to genomic data. However, continued progress and new advances will require expanded collaborations. Formal efforts to bring these groups together through, for example, workshops, shared funding, and other opportunities would reap large rewards for the design and extension of cyberinfrastructure in capturing the many elements of the extended specimen and aligning data resources in living and natural history collections.

CONCLUSIONS

Certain impediments will have to be overcome before the potential of a national cyberinfrastructure and the digitization it supports can be realized. Through varied programs past and present, NSF contributions to biological collections digitization and cyberinfrastructure have been critical in the United States. To be successful and sustainable, the digitization and development of a robust cyberinfrastructure will require continued support from NSF. Although digitization efforts have involved hundreds of collections, phylogenetic, geographic, temporal, and taxonomic gaps in digitization are evident. Harnessing the opportunity for data-driven discoveries and transdisciplinary collaboration will depend on a continuing effort to digitize new and existing biological collections using developed communities of practice (e.g., best practices and standards). Investment in the development of new technologies and cost-effective, high-throughput workflows for digitizing collections that, to date, have lagged—such as entomological collections—will enhance both the number of specimens and the taxonomic scope of digitized collections. Future digitization initiatives will need to be prioritized to address this disparity in order to ensure better representation of data from these underrepresented groups. In addition, the identification, assessment, and accessioning of legacy project-based collections could bring a large number of valuable specimens and their digitized records into the public domain and prevent the future accumulation of inaccessible collections that diminish NSF's investment in their assembly and future use in research and education. Compounding these issues is the lack of resources or associated workforce (see Chapter 6) and also staff who may not realize the value of the collections once digitized. If these dark data can be made available, both the physical collections and their digital representations can be used in future research, contributing to the growing fabric of networked collections.

National and global portals and catalogs have made important contributions to the biological collections community by providing a platform with which to exchange and share data and promote standardization and consistency. Continual updating, augmenting, and improving digital data records using annotation tools and data assertions, for example, will greatly improve overall data quality and, in turn, lead to more comprehensive data integration and greater accessibility of digital data. However, mechanisms for data annotation and attribution require an interoperability of data and systems, which may be impeded by global indecision about the application of globally unique identifiers for specimen records. In addition, despite some progress, integrated systems that enable the citation of data used in research publications and attribution to data providers are difficult to develop and will require an all-encompassing approach with social incentives and innovative technological solutions. These are not insurmountable problems, but it will be important to address them in the development of a comprehensive national cyberinfrastructure for the large-scale, long-term digitization and use of digitized data.

The integration of specimen data with other biological components as well as with data sources outside of the biological realm will require the implementation of a network of cyberinfrastructure resources not yet realized. Possible future collaborations are potentially unlimited, but computer scientists and the collections community will require mechanisms to bring them together and instruction on how to communicate across disciplinary barriers. Rapid developments during the past few years argue for the value of these collaborations. Just as innovations in digitization have resulted from partnerships between these communities, further collaborations, particularly in the application of machine learning, will lead to even greater progress in digitization, georeferencing, and data analysis. A unified cyberinfrastructure that connects all types of biological collections, such as living and natural history collections, could accelerate research and provide innovative educational opportunities. Moreover, a permanent national cyberinfrastructure that supports the needs noted above in terms of expanded digitization of dark data, improvement in data quality, and an increased accessibility to digital data would certainly spur data use. Without this resource, collections—both physical and digital—will continue to be underused.

RECOMMENDATIONS FOR THE NEXT STEPS

Recommendation 5-1: The leadership (managers and directors) of biological collections should provide the necessary mechanisms for staff to keep pace with advances in digitization and data management through training in digitization techniques and publishing of standardized quality data that can be efficiently integrated into portals.

Recommendation 5-2: Professional societies should initiate and cultivate opportunities for research collaborations within the biological collections community. These collaborations should include working with the computer and data science communities to promote the development and implementation of tools to build the cyberinfrastructure (e.g., data storage, annotation, integration, and accessibility to expand the use of biological collections to a broader range of stakeholders).

Recommendation 5-3: The National Science Foundation (NSF) Directorate for Biological Sciences should continue to provide funding for the digitization of biological collections and for the cyberinfrastructure to support both living and natural history collections. Specifically, the NSF Directorate for Biological Sciences should:

- partner with other directorates within NSF (e.g., physics, chemistry, computer science, and education) and other federal agencies and departments (e.g., the Department of Health and Human Services, the Department of Agriculture, the Food and Drug Administration, the Department of the Interior, the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration, the Department of Energy, etc.);
- establish ongoing mechanisms for the biological collections community to meet, develop best practices, and work toward goals such as establishing and implementing unique identifiers, clear workflows, and standardized data pipelines; and
- promote and fund the development of a necessary national cyberinfrastructure, with appropriate tools, and technology to affect the efficient multi-layer integration of data and collections attribution.

6
Cultivating a Highly Skilled Workforce

BIOLOGICAL COLLECTIONS REQUIRE PERSONNEL WITH MULTIFACETED AND COMPLEX COMPETENCIES

Effective leadership is a critical factor in research quality and integrity, as well as long-term sustainability of infrastructure (Antes et al., 2016; NRC, 2005, 2014b). Curatorial and technical skills can enable biological collections to serve as a nexus for transdisciplinary research by ensuring that specimens and data are accessible to the broadest range of users, including scientists, educators, students, entrepreneurs, and others. Teaching and public communication skills can stimulate curiosity and engagement across a wide range of learners and stakeholders. Despite the importance of these skills, the biological collections *workforce pipeline*—a conduit that extends from when a future collections professional first becomes aware of biological collections to established professionals seeking to enhance or learn new skills to near-retirement collections professionals for whom it is imperative to transfer their knowledge to other staff or new hires—is underdeveloped. This chapter describes the primary challenges and opportunities to understand, build, and support a thriving, diverse workforce ecosystem for biological collections.

THE BIOLOGICAL COLLECTIONS WORKFORCE ECOSYSTEM

Highly skilled, trained personnel underlie the increasing sophistication in the ways that biological collections carry out their missions and meet the dynamic needs of science, education, and broader society. The responsibilities associated with leading a vibrant biological collection are similar to and as complex as those required to maintain any research center or innovation hub. The diversity of the nation's biological collections contributes to the breadth and depth of this range of expertise.

Staffing models vary among biological collections due to differences in institutional missions, size, diversity of taxa, and financial support, but can involve a combination of a director, curators, and collection managers. For example, the Department of Ornithology in the Biodiversity Institute & Natural History Museum at The University of Kansas houses 123,000 bird specimens, and employs two Ph.D.-level curators and one Ph.D.-level collection manager. In contrast, the ornithology department at the Museum of Comparative Zoology at Harvard University, which houses ~400,000 specimens, one of the largest collections of bird specimens in the world, employs only one Ph.D.-level curator and 2.5 full-time equivalent staff positions. The lack of a set staffing model is similar to other distributed research infrastructures such as field stations and marine laboratories, which also vary in size and complexity (NRC, 2014b). Nonetheless, the basic responsibilities of biological collections personnel span a continuum across three broad categories: (1) leadership and management, (2) science and technology, and (3) teaching and public engagement.

Leadership and Management

The responsibility for leadership and management of a biological collection varies among different institutions. Based on the committee's extensive experience, the leadership and management of a biological collection are carried out by one person in some instances, usually an institutional director, a curator, or a collection manager. In other cases, leadership responsibilities are divided among two or more people according to a hierarchy. In general, an institution director sets a vision for the institution, of which biological collections are one part. Collection managers usually oversee day-to-day collection maintenance and can also contribute to developing and implementing strategies for collections growth as well as advocate for resources for long-term sustainability. Curators are chiefly responsible for establishing a vision for their collection and then setting the direction of growth, use, and ultimately long-term sustainability of their collection. Therefore, the success and impact of a particular collection are often closely tied to the direct engagement and oversight of the curator. For this reason, hiring decisions for leadership of a biological collection are often based on academic expertise, familiarity with the taxonomy and history of the collection, and administrative expertise (Krishtalka and Humphrey, 2000).

Effective leaders are responsible for building, providing, and maintaining the infrastructure that enables research to thrive, as well as providing the vision and guidance to lead an organization forward (Hao and Yazdanifard, 2015). Leadership and management require skills that include developing and implementing strategic plans and business strategies; fundraising; financial, personnel, and information management; regulatory compliance; entrepreneurship; evaluating efficacy and impact; and communications. Because a key feature of biological collections is to distribute or loan specimens and associated biological research materials, the personnel must also employ the same skills in customer (i.e., user) engagement and support as needed in libraries and other public service organizations. Biological collections leaders navigate complex national and international law, meet requirements of biosafety and security, and evaluate, articulate, and enhance scientific impacts of their collections. Leadership and management encompass planning for the space to accommodate expansion through curation and future acquisitions. Predicting the space, expertise, and number of personnel needed for collections growth is a critical task necessary to ensure that specimens and their data remain well cared for and available well beyond the length of a career.

Leaders also set the organizational culture and cultivate durable relationships with employees, students, funders, host institutions, and members of the public (NRC, 2014b, 2015a). Therefore, the long-term sustainability of biological collections requires leaders who not only are scholars, but also hold skillsets found in executive directors, entrepreneurs, research coordinators, government and regulatory affairs coordinators, database managers, and development and public affairs officers.

Scientific and Technical Staff

Science and technology are integral to a biological collection. Biological collections need personnel with specialized expertise to curate and care for specimens and specimen data. Biological collections personnel usually hold postgraduate degrees in the field of biology most related to the collection. A museum studies degree can also lead to a career in natural history collections management. The level of education needed for different collections positions can vary depending on the history and traditions of different institutions. A Ph.D. in a relevant scientific discipline is required for collection curators and often collection managers at some institutions, but others employ managers with an undergraduate or master's degree (Bakker et al., 2020; Pennington et al., 2013; Shi et al., 2011; Thiers et al., 2019; Wu et al., 2017). Many collections personnel are scientists, and often lead or collaborate on biological collections–based research in addition to upholding their responsibilities to curate and distribute specimens and data.

By collecting, maintaining, and generating specimens and their associated data appropriately, trained personnel ensure the utility of biological collections for a variety of research and educational purposes as well as make it possible for biological collections to adjust more nimbly to a wider set of scientific purposes than those considered when the collection was originally assembled. Such new uses of biological collections and their data are expanding not only in life science research, but also in the physical sciences, mathematics, engineering, geography, arts, and other fields (Heberling and Isaac, 2017; Schindel and Cook, 2018). In living stock collections, scientific personnel are responsible for ensuring a strong platform for reproducibility and replicability[1] in research, two hallmarks of scientific rigor and validity (McCluskey, 2017; NASEM, 2019d).

Teaching and Public Engagement

The breadth of teaching and public engagement activities varies among collections. Many biological collections personnel teach and mentor students and postdocs, as well as participate in public outreach activities such as tours, exhibits, and virtual public programs. This is evidenced by the rich diversity of collections-based formal educational programs and research experiences

[1] The National Academies report *Reproducibility and Replicability in Science* (NASEM, 2019d) defines reproducibility as obtaining consistent computational results using the same input data, computational steps, methods, code, and conditions of analysis. Replicability is defined as obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data.

for students of all grade levels, informal education programs, and a variety of collections-focused public engagement opportunities and activities, including citizen science programs (see Chapter 3). The value of biological collections for education is so significant that some members of the biological collections community have issued calls to use biological collections as a foundational teaching tool in science, technology, engineering, and mathematics (STEM) (Cook et al., 2014; Monfils et al., 2017; Powers et al., 2014). Powers et al. (2014) also outlined grand challenges for natural history collections that could be expanded to living collections and serve as organizing principles for an education-focused community of practice (see Box 6-1).

Also highlighted in Chapter 3, collections-based education, training, and public outreach programs can be used to increase participation of historically underrepresented groups in STEM (Miller et al., 2020b). These activities benefit natural history collections via augmenting collections and increasing public investment while simultaneously benefiting participating communities through increased knowledge, transparency, and involvement in research and community-relevant decision making (Ballard et al., 2017; Haywood, 2014; Roger and Klistorner, 2016).

Volunteers, Student Interns, and Postdoctoral Fellows

To meet some of their scientific, technical, educational, and public engagement needs, some biological collections employ part- or short-term staff and undergraduate and graduate student interns or recruit volunteers. Students, volunteers, and docents play a large role particularly in natural history collections and living biodiversity research collections,[2] especially with preparing and digitizing specimens and processing specimen loans. Volunteers at the New York Botanical Garden (NYBG), for example, account for one-quarter of the workforce needed to prepare specimens and deposit them into collections (Barbara Thiers, NYBG, personal communication, 2019). Volunteers at NYBG also image ~100,000 specimens (new accessions or as part of retroactive imaging projects) and transcribe data for an average of 50,000 specimens per year. In some cases, volunteers also participate in identifying and curating specimens. Although the involvement of volunteers in collections is highly desirable as a means of public outreach and providing the best possible collections care, their contributions to collections work may mask the inadequacy of the institutional budget for collections personnel.

Postdoctoral fellows also contribute to collections curation and management, but typically as part of the research project that funds the postdoctoral

[2] Living biodiversity collections that emphasize education and public outreach, such as the Duke Lemur Center, often involve volunteers in their work. By contrast, very few living stock collections use volunteers or involve citizen scientists for collections management. This is, in part, because maintaining living stock collections requires advanced disciplinary education and expertise to maintain the genetic integrity of the specimens, and also because of liability issues related to the biosafety of the materials involved.

BOX 6-1
Considerations for a Biological Collections–Focused Community of Practice in Science, Technology, Engineering, and Mathematics Education and Workforce Development

In 2014, scientists and experts from seven institutions across the United States published a vision and strategy to revolutionize the use of biological collections in education (Powers et al., 2014). The authors emphasized that bringing scientists and educators together opens opportunities for the use of biological collections in education at all levels and throughout life. They reasoned that the primary mechanism and grand challenge to realize those opportunities is through the integration of the following three resources:

1. Specimens and Collections. Scientific specimens are a national resource for engaging people of all ages, stimulating inquiry about past and present life on Earth.
2. Specimen-Based Electronic Resources. Electronic resources expand access to a greater range of information and expand the possible types of specimen-based inquiry (e.g., exploring local biodiversity) and skill development (e.g., data mining).
3. People and Human Resources. Interactions among collections-based scientists, educators, and students build bridges between disciplinary silos and create a more inclusive scientific enterprise.

The authors focused primarily on the use of natural history collections for kindergarten through undergraduate students' formal and informal education. However, they articulated a grand challenge that could potentially be broadened for use in all biological collection types for a collections-based community of practice for lifelong learning and workforce skill development in science, technology, engineering, and mathematics.

position rather than through formal collections training. The National Science Foundation's (NSF's) Postdoctoral Research Fellowships in Biology[3] have been instrumental in exposing early-career scientists to biological collection careers. This 5-year program, which made its final awards in 2020, has provided fellowships for creative research using biological collections to more than 80 early-career scientists. Although the program does not involve formal training in collections management, fellows obtained experience in the use and organization of collections and collections data.

[3] See https://www.nsf.gov/publications/pub_summ.jsp?WT.z_pims_id=503622&ods_key=nsf19597.

Education and Professional Development of Biological Collections Staff

Workforce education, training, and professional development are investments in knowledge, skills, and abilities to ensure a sustainable, productive research infrastructure. However, a collections workforce development pipeline, which begins with a college education and leads to a biological collections career (with professional development opportunities), is somewhat informal. Consequently, there is still a lot unknown about the biological collection workforce including its size, demographics, the scope of responsibilities, and systems for professional recognition and rewards. There has not yet been a robust analysis to identify gaps in education and professional development opportunities for critical skillsets, or of the efficacy of available education and training mechanisms to cultivate a diverse and inclusive workforce ecosystem as well as address training needs as new technologies and challenges arise. Community colleges are well suited for targeted purpose-driven STEM education and they reach a diverse student body (NRC, 2012b).

Most collections personnel are currently trained on the job. Relevant scientific training (e.g., specimen collection, verification, preparation, curation, and maintenance) can take place in biology and natural history courses on specific organismal groups (e.g., microbiota, insects, fish, and birds), although these types of courses are in decline nationally (Hiller et al., 2017; Scott et al., 2012; Tewksbury et al., 2014). A small number of museum studies programs offer formal degree or certificate programs for natural history collection work. Collections knowledge can also be passed down from curators who serve as advisors to students (Leather and Quicke, 2009).

Some professional societies offer training and professional development through topical workshops and conferences. The entomology community, for example, occasionally offers a collections management workshop,[4] and in recent years, the Society of Herbarium Curators has offered in-person and online training in strategic planning.[5] The Society for the Preservation of Natural History Collections (SPNHC) regularly offers workshops and training in new aspects of permit and collections compliance as well as exposure to advanced topics in curation and conservation. Many professional societies associated with living stock collections, such as the World Federation for Culture Collections (WFCC) and the American Society for Microbiology, also offer professional training programs. Because poor product quality can harm human health and the environment, biobanks and biological resource centers have technical education and certification programs for living stock collections (OECD, 2004).

[4] See https://ecnweb.net/workshop.
[5] See http://www.herbariumcurators.org/strategic-planning-course.

CHALLENGES

Biological collections face a number of complex and interrelated workforce challenges. Most biological collections are understaffed. Existing staff shortages can slow the pace of curation, digitization efforts, and distribution of specimen loans and reduce visitor resources even as more researchers are using collections (Schindel and Cook, 2018). Recruiting exceptional leaders to not only manage a biological collection but also lead biological collections into the future with an eye toward innovative collecting, curating, and research is another significant challenge. Cultivating a highly skilled, well-trained, and diverse biological collections workforce also requires attention to several intersecting issues: insufficient number and diversity of trained staff; the limited availability of relevant academic pathways to foster the next generation of the biological collections workforce; and inadequate coordination among existing training and professional development programs to enrich and expand the skillsets and diversity of the current biological collections workforce and leadership. Underlying all of these challenges is the need for consistent and collaborative mechanisms to monitor workforce trends in order to better identify and strategically address needs and gaps among the nation's biological collections ecosystem.

Differing organizational structures, institutional cultures, and systems of compensation and professional recognition for the range of responsibilities and outstanding performance create additional layers of complexity to workforce challenges. It is beyond the scope of this report to delve into the additional complicating factors, but they will be important for biological collections leadership and NSF to bear in mind as they grapple with workforce challenges. This section focuses on the challenges that most impact the availability and preparedness of personnel with the required expertise for the use and maintenance of collections for research and education.

Insufficient Number of Trained Staff in an Environment with a Multifaceted and Expanding Range of Necessary Skills

The increasing sophistication and global nature of science, the financial demands that accompany ongoing maintenance of research infrastructure, and entreaties by collections stakeholders and funders for more innovation and accountability are expanding the responsibilities and expectations of the limited number of biological collections personnel. In addition, changing needs within the biological collections field are also placing new demands on biological collections leadership and staff. For example, there are calls for a new type of curator, whose responsibilities would combine research with active and in-depth public engagement, working collaboratively with educators, and utilizing social media as a means for frequent and ongoing public communication (Dance, 2017; Jarreau et al., 2019; Lessard et al., 2017). The need to juggle competing priorities is exacerbated when biological collections staff absorb the increased workload created when managers lack authority and funding to fill vacancies due to reductions in force, including retirement without replacement (Miller et al., 2020b, p. 3). Prioritizing sufficient time for leadership, scientific, technical, and teaching responsibilities at larger, well-funded biological collections is difficult, and even more so at smaller collections that may not have the support to hire new staff for these very different functions.

The breadth of important skills and responsibilities, from strategic leadership to curation and care of specimens to coordinating access to digital information, has led a number of professional organizations to provide some guidance about staffing (e.g., AIC, 2013; OECD, 2007). Based on these guidelines, Smith et al. (2014) suggest that biological resource centers employ six full-time staff to meet the curation, quality control, order fulfillment, and regulatory compliance needs of a modest-size living stock collection of 5,000 to 10,000 specimens that distribute 2,000 specimens per year (Smith et al., 2014) (see Table 6-1). The guidelines may not capture the variability in staffing needs or capabilities of different institutions, which today often reflect traditions of hiring and growth rather than specific collection needs. Nevertheless, the robust staffing levels recommended by Smith et al. (2014) are uncommon for biological collections in the United States. For example, of the 22 active U.S. culture collections registered in the World Data Centre for Microorganisms (Wu et al., 2017),[6] only 3 employ 6 or more staff members.

The Biological Collections Workforce Pipeline Is Underdeveloped

The biological collections community lacks a formal and clearly defined workforce pipeline—one that takes into consideration education and training needs before, during, and as staff transition into a collections career. The lack of an efficient and robust workforce pipeline inhibits the ability of biological collections leaders as well as the biological collections community to anticipate and strategically plan for cultivating a robust workforce. Three particular parts of the workforce pipeline with significant challenges are (1) cultivating the next-generation biological collections leadership, scientific, technical, and education staff; (2) coordinating professional development opportunities for the existing collections workforce as new skills, technologies, and challenges arise; and (3) developing a more diverse professional workforce. A conceptual paradigm shift is needed to enhance education and workforce development for the long-term sustainability of biological collections. This section describes key major obstacles that impede such a paradigm shift.

Insufficient Education Programs to Support the Development of the Next-Generation Biological Collections Workforce

There are few university degrees or certificate programs that include collections-focused curricula. Of the 185 museum studies programs in the United States, only 25 (e.g., University of Colorado Boulder, Texas Tech University,

[6] The World Data Centre for Microorganisms is a global registry for WFCC. See http://www.wdcm.org.

TABLE 6-1 Recommended Staffing Guidelines for a Microbial Biological Resource Center with 5,000 Strains

Staffing Responsibilities	Minimum Recommended Number of Staff [a]
Collection management and business development	1
Authentication, preservation, and distribution of microbial strains	3
Implementing quality standards and adherence to regulations	1
Identification services	1
Total	6

[a] Additional staff may be required depending on taxonomic diversity (depth and breadth), desired research capacity, and other services provided by the biological resource center.
SOURCE: Smith et al., 2014.

The University of New Mexico, The University of Kansas, and the University of Florida) offer a specialized focus on natural history collections.[7] These museum programs focus primarily on scientific and technical aspects of natural history collections management and are not designed for teaching about the management or curation of living stock collections.

Further complicating the educational landscape, the breadth of expertise required to manage biological collections is changing. Although taxonomic expertise and general collections best practices may have been sufficient in the past, there is now a need for education and training in strategic leadership and business management (see above and Chapter 7), data science (see Chapter 5), and new scientific methods and advanced technologies in order to address expanding research missions of many modern biological collections. The need to cultivate future biological collections leaders reflects repeated calls throughout the scientific community for scientists to receive leadership education and training (Kvaskoff and McKay, 2014; Leiserson and McVinny, 2015). As noted in the National Research Council report *Enhancing the Effectiveness of Team Science*, there are more than 50 years of research on organizational leadership that provides "a robust foundation of evidence to guide professional development for leaders of science teams and larger groups" (NRC, 2015a, p. 146). However, leadership training for scientists is not widely integrated into university curricula, nor are there sufficient and consistent professional development opportunities for established scientists.

Training and Professional Development Options for the Existing Workforce Are Uncoordinated

The available collections-focused education and training opportunities are insufficient to grow and support a robust workforce pipeline. In addition, few of

[7] See http://ww2.aam-us.org/resources/careers/museum-studies.

these efforts are well coordinated. There are minimal mechanisms for collaboration, and no guidelines or standards that ensure the quality and consistency of curricula among the existing education and training opportunities. The lack of a consistent and structured mechanism for workforce development can impact day-to-day operations; the speed at which advanced methods, such as digitization, are adopted; and the development of innovative scientific or educational uses for specimens. This situation also leads to incomplete knowledge about the history of the collection and best practices to maintain it, past preservation techniques (especially with regard to the use of hazardous materials), legacy data products, archives, dates of acquisition of major equipment and service agreements, and the breadth and depth of stakeholders including local regulators having jurisdiction over the collection, safety officials, funders, volunteers, and others. A loss of this critical knowledge is especially high risk if there is a gap between the departure of one collections curator or manager and the hiring of the next.

Limited Efforts to Broaden Participation in the Biological Collections Workforce

The lack of a formalized workforce pathway to biological collections careers is a limiting factor in efforts to develop a more diverse professional workforce. The 2011 National Academy of Sciences, National Academy of Engineering, and Institute of Medicine report *Expanding Underrepresented Minority Participation: America's Science and Technology Talent at the Crossroads* raised an alarm that the STEM education and workforce are seriously lacking participation by individuals from historically underrepresented communities (NAS et al., 2011). Since 1990 the STEM workforce has nearly doubled (9.7 million to 17.3 million), and Black and Hispanic workers continue to be underrepresented (Pew Research Center, 2018). Although doctorates are not required for all job positions, they can be telling indicators of diversity and inclusion in STEM. In the field of geoscience, a collections-related discipline, Bernard and Cooperdock (2018) indicate little improvement in the diversity of doctorates in the United States over the past four decades, despite outreach efforts aimed at shifting the demographics. In 2016 minority groups comprised 31 percent of the U.S. population, yet received only 6 percent of geoscience doctorates awarded to U.S. citizens and permanent residents—the lowest proportion of all STEM fields. More than 87 percent of respondents to an online survey of faculty associated with ecology and evolutionary biology doctoral programs in the United States, another collections-associated field, identified as White/Caucasian (Jimenez et al., 2019). Dutt (2020) further emphasizes that progress toward diversification can only come with a concerted shift in mindsets and a deeper understanding of the complexities of race. The 2020 National Academies report *A Vision for NSF Earth Sciences 2020–2030: Earth in Time* reiterates these findings and recommends a more significant investment from the NSF Division of Earth Sciences on issues of diversity, equity, and inclusion within the field (NASEM, 2020).

Specimen-based education and training are essential to introducing a broader range of scientists to the potential of a collections career. Biological collections provide an opportunity for students and postdoctoral fellows to observe and experience a wide variety of potential career options that might include research, education, or collection-specific careers. For example, NSF's Integrated Digitized Biocollections (iDigBio) national resource has held workshops specifically to address broadening participation in the biological sciences with the goal of introducing students, especially those in underrepresented populations, to museum and biodiversity science careers.[8] In particular, community colleges are well distributed throughout the United States and are well suited for reaching a more diverse audience and preparation of the STEM workforce (NRC, 2012b). However, the lack of focused efforts to recruit, support, and retain a diverse professional workforce at all stages of the workforce pipeline constrains current efforts to cultivate a more diverse future workforce.

Insufficient Institutional Recognition and Support for Collections Curation and Care

Few professional mechanisms provide guidance, training, and professional recognition for curatorial work, yet the success and impact of a biological collection are closely tied to the direct engagement and oversight of a curator. Nationwide reports of decreasing institutional support for biological collections and their associated curatorial staff and resources parallel trends of biological collections being closed or transferred (Dalton, 2003; Gropp, 2003, 2004; Schmidly, 2005; Winker, 2008). For example, within the past 20 years, 45 mammal collections in the Western Hemisphere, approximately 10 percent of the total number, were closed or transferred (Dunnum et al., 2018). Thirty-one of these mammal collections were held by U.S. universities. Reinforcing this trend is the lack of widespread recognition of how biological collections contribute to science and society generally, and to an individual institution's mission and reputation specifically, which results in hiring priorities and funding initiatives that lack an explicit focus on building a robust infrastructure and workforce (Schmidly, 2005).

At universities and some large museums, biological collection curators are often tenure-track positions. As a result, decisions about hiring and job advancement typically focus on an individual's research, teaching, and public service accomplishments—the three hallmarks of academic tenure and promotion. Professional recognition, compensation, and performance review of curators often do not explicitly detail, evaluate, or incentivize curatorial responsibilities, even if they intersect with research, teaching, and public service. In addition, curators are not always recognized for their leadership responsibilities, which include long-term planning, commitment, and administration of the physical space and intellectual capital of the biological collection, and the ways in which the collection contributes to the reputation and standing of the institution as a

[8] See https://www.idigbio.org/tags/broadening-participation.

whole. These duties are similar to that of academic department chairs, deans, and the heads of research institutes. Hence, a valuable component of the U.S. research portfolio, curation of biological collections, is inadequately incentivized and supported.

THE WAY FORWARD

The long-term sustainability of the nation's biological collections will require deliberate action for an even and equitable development of the workforce. A renewed emphasis on education, training, students, and staff is essential for the continued success and future of biological collections (Miller et al., 2020b). To be effective, the future biological collections workforce will require innovative and comprehensive approaches to identify and address their needs. This section describes important first steps forward.

Launch a Community-Wide Conversation on Critical Skillsets

The ability to develop adequate and consistent education and training programs depends, in part, on identifying the critical, broadly applicable skills needed to manage a biological collection and promote and expand its use for research, education, and other purposes. The focus of the available curricula is on the scientific and technical skillsets. There is a considerable wealth of collective knowledge among the many professional societies and networks, such as WFCC, Natural Science Collections Alliance, SPNHC, and iDigBio, in regard to those skillsets. However, there has not yet been a community-wide conversation about other critical skillsets, such as leadership, business management, informal science education, public communication, and impact evaluation, for which there are no consistent, collection-focused education or training programs.

A parallel situation prompted the cyberinfrastructure facilitation community[9] to identify critical skillsets and then launch a series of NSF-supported Virtual Residency workshops to build capacity (Neeman et al., 2018). The 2018 workshop focus areas are similar to or the same as desired skillsets discussed among many biological collections: leadership, expertise in rapidly changing technology, funding acquisition, outreach, and communication. This Virtual Residency workshop was able to reach 216 participants from 147 institutions across 42 states, 2 U.S. territories, and 2 other countries.

A comprehensive list of the range of skills needed for successful collections management could inform a collaborative effort among the nation's biological collections to outline roles and responsibilities of biological collections directors, curators, managers, and other positions, and clarify appropriate career pathways. This in turn could inform existing university programs or efforts to develop new collections-focused curricula. It could also incentivize professional societies to circulate information about available learning resources that would

[9] Cyberinfrastructure facilitators work closely with scientists to help them use research community systems and services.

help members of their community achieve accreditation. This information might increase the number of collections-focused education and training programs or better calibrate existing ones to workforce needs and the skills most critical to a successful biological collections career.

Monitor and Evaluate Workforce Capabilities and Needs

The ability to identify, monitor, and evaluate progress will require mechanisms to collect and analyze workforce data. Issuing a periodic survey is a valuable mechanism that many professional communities use to collect workforce data. For example, in 2004 the National Association of Social Workers (NASW) conducted a comprehensive, benchmark survey that included questions about their workplace, professional responsibilities, demographics, education and training, compensation and benefits, and even their perceptions about the social work profession (Center for Health Workforce Studies and NASW Center for Workforce Studies, 2006). NASW used the survey data to inform policy and planning decisions to cultivate a social work workforce that is well equipped for the needs of the nation. In 2013, the National Association of Marine Laboratories (NAML) and the Organization of Biological Field Stations (OBFS) joined forces to survey their community (NAML and OBFS, 2013) about infrastructure capabilities and needs, including staffing models. NAML and OBFS used the survey data, a community-wide workshop, and feedback from researchers and other stakeholders who use field stations and marine labs to inform the development of a national strategic vision (Billick et al., 2013). WFCC and the Biodiversity Collections Network have both invested in survey mechanisms to identify the specific needs of their respective communities, although the workforce was not the primary focus of those endeavors. Pooling data about the nation's biological collections workforce could enable a strength, weakness, opportunity, and threats analysis (Gürel and Tat, 2017) as a way to monitor workforce trends and assess the effectiveness of workforce development strategies. This type of analysis would facilitate interactions between the biological collection community and relevant professional communities, and also facilitate a community-wide conversation identifying critical skillsets and strengthening the biological collections workforce pipeline. Starting with recruitment and training of new staff, the discussions about the workforce pipeline will also need to address retention, re-skilling existing staff, succession management, and integrating volunteers.

Promote Diversity as an Integral Element of the Workforce Development Pipeline

The many complex problems addressed by the biological collections community require an innovative workforce, with broad and varied backgrounds. Increased diversity benefits scientific advancement: different perspectives and experiences spark novel questions, improve problem solving, enhance the effectiveness of teams, and generate higher-impact science (Disis and Slattery, 2010; Freeman and Huang, 2015; Medin and Lee, 2012; Nielsen et al., 2017; Valantine and Collins, 2015). Collaborations that involve a diverse group of people are more likely to tackle problems in creative ways that can lead to higher levels of scientific innovation (Hong and Page, 2004). Campbell et al. (2013) found that increasing the diversity of biological collections staff, and those using those collections, is likely to have benefits for both the users and the collections.

Promoting a more diverse workforce needs to be an integral aspect of discussions about the biological collections workforce pipeline (Papers that matter, 2020). Rethinking traditional models and paradigms for how biology and paleontology are taught is a critical first step toward increasing diversity in these fields (Visaggi, 2020) and in turn, biological collections specifically, because most of the workforce is trained through these disciplines. Course-based undergraduate research experiences (CUREs) provide authentic research training that can reach students who might not be able to afford to volunteer. The Biodiversity Literacy in Undergraduate Education and the Biodiversity Collections in Ecology and Evolution Network are examples of programs implementing collections-focused CUREs. Paid internships and mentoring opportunities that are dedicated to increasing participation from traditionally underrepresented groups in science are also beneficial strategies. This also might mean rethinking and restructuring how such opportunities are advertised and offered so that they are more visible and accessible to underrepresented students. Professional societies have an important role in this work (e.g., establishing committees or working groups on diversity and inclusion; addressing discrimination, harassment, and bullying in codes of conduct and ethics; promoting the work of diverse members; mentorship programs and funding for students to attend professional meetings), and some already do so. For example, in 2015 the American Elasmobranch Society established the Young Professional Recruitment Fund[10] diversity scholarship, a competitive award for individuals from historically underrepresented groups in marine science or who are performing research in a developing nation, which provides professional development training, mentorship, and a 1-year membership. The American Society of Plant Taxonomists also funds an early-career research grant to support the professional development and retention of botanists from underrepresented groups. The Paleontological Society occasionally offers Conference Travel Grants to Support Inclusion[11] Inclusion that are competitive grants to offset the travel costs of members from underrepresented or at-risk groups who otherwise could not attend the Geological Society of America annual meeting. Community colleges are a potentially strong component of training a diverse and STEM-literate workforce (NRC, 2012b) and while they are supported by some federal programs (e.g., the Department of Agriculture's Community Facilities Program), additional support from other federal agencies or other funding sources would be useful to target the biological collections workforce pipeline.

[10] See https://elasmo.org/young-professional-recruitment-fund.
[11] See https://paleo.memberclicks.net/grants-and-programs.

Harmonize Available Staff Training and Professional Development Opportunities

Professional societies could provide the expertise and collective engagement needed to identify needs and deliver training in a variety of forms. Although biological collections vary in the types of specimens and materials they curate, certain aspects of curation and management extend across all collections. Professional societies could work together to pinpoint those common elements—which might include the use of standards in data management, best practices for databasing, interpreting and implementing best practices related to the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity, and more—and then collectively consider mechanisms for joint training. For example, the American Institute of Biological Sciences, which is an umbrella organization for other professional biological societies, regularly offers short courses and "bootcamps" on topics of interest to its member organizations and also provides training for professional leaders on building resilient scientific societies. Training opportunities such as these are valuable in themselves for the participants, but they also provide opportunities for attendees from different organizations to learn from each other. This model of professional societies working together could be extended to the development of joint certification programs that target key needs for collections personnel. Networks of collections may also fill this collaborative role. For example, the Microbial Resource Research Infrastructure is a pan-European network of more than 50 biorepositories that has collectively developed best practices and training guidelines to ensure certification of its member collections.

Collaborations among collections—whether through professional organizations or networks—could draw on collective expertise nationally without putting a huge burden on any single community and could yield well-trained collections professionals across disciplines. Ideally, a shared central resource announcing training events such as workshops and hosting online materials could promote a greater sense of community among collections, provide collective best practices, and allow access to professional development to all collections, regardless of size and financial status. As with other responsibilities of biological collections leaders and staff, the challenge for teaching and public engagement will be to remain agile and responsive to current conditions and needs of the respective communities. The current coronavirus disease 2019 crisis has clearly shown that organizations equipped to respond with web-based materials, lessons, and other means of online engagement have facilitated their reach and utility.

Moreover, professional organizations have the ability to attract new members and young professionals to their fields, so coordinated offerings on the value of the national (and global) collections infrastructure and its many uses could be important for filling the pipeline with the next-generation workforce. A national message could elevate the vision of collections personnel as part of an interconnected national scientific infrastructure.

Connect with Relevant Communities of Experts

The biological collections community does not need to reinvent the wheel to find ways to develop and structure all education and professional development opportunities. As the development of the biological collections workforce increasingly demands transdisciplinary skills, it is important to recognize opportunities to partner with professional communities and use their resources to supplement skillsets. Much could be learned from disciplines with established formal programs and training modules that parallel the needs of the biological collections community. For example, library science, the study of collecting, preserving, cataloging, and making documents available in libraries, could be an important source of evidence-based practices and guidance on the dynamic nature of information management. Museum studies, archival science, and more recently data science are other disciplines with pedagogy that parallels the workforce training needs of the biological collections staff. It might also be feasible to partner with business schools and their nonprofit management programs to develop some focused coursework or a certificate program specific for biological collections and the challenges inherent to them. Similarly, biological collections experts could connect to a range of educators (in informal or formal settings) and work in partnership with them to reach students at all levels and lifelong learners. All of these learners have scientific interests, questions, and needs (Bakker et al., 2020; Pennington et al., 2013; Shi et al., 2011; Thiers et al., 2019; Wu et al., 2018). There may be educational staff within host institutions with whom biological collections staff could collaborate.

Innovate Staffing Strategies

Developing a robust workforce pipeline and the resources to support a greater number of staff will take time. Meanwhile, biological collections need near-term solutions to staffing shortages. What might be some immediate approaches that do not rely as heavily on complex workforce analyses or substantial increases in funding? Formalizing volunteer and citizen science efforts and integrating these contributions as a means of filling staffing shortages or needs is one possibility. Metrics to track and monitor these efforts, their effect on workforce needs and capabilities, and their potential impact on existing staff time and budgets are important. A possible additional benefit would be in terms of public engagement and education. While citizen science is and will continue to be a valuable contributor to biological collections (McKinley et al., 2017), the role of citizen science has increased with the digitization of some biological collections, and we can clearly see how citizen science can directly impact specimen-based researchers by facilitating the digitization process (Ellwood et al., 2015, 2018).

Integrating student research and internship opportunities with curation or other aspects of collections maintenance is another possible approach. Such integration could help build awareness of and support for biological collections, and provide an important path to ensuring that the collections workforce is maintained and enriched over time. For example, the Museum of Vertebrate

Zoology (MVZ) at the University of California, Berkeley, runs the MVZ Undergraduate Program,[12] which has allowed hundreds of undergraduates to become involved in vital and impactful activities in the museum's collections in exchange for course credit (Hiller et al., 2017). The program at the University of California, Berkeley, may be a model for other universities with biological collections. It combines training in collections care with essential research skills and exposure to biodiversity not available elsewhere in the university. The students enrolled in the program (more than 100 per semester) also provide much needed workforce support to the collections, creating a mutual benefit. Use of a tiered structure, wherein students master one set of curatorial techniques before advancing to higher-level work, increases interest for highly motivated students, and encourages retention. Students at the highest level help supervise the beginners, may present independent research at conferences, and receive academic credit for their participation (Hiller et al., 2017). Similarly, the Biodiversity Institute & Natural History Museum at The University of Kansas has strong ties with the Museum Studies degree program on campus and provides valuable internship[13] (degree requirement) opportunities for its many students, who in turn provide much needed assistance in the collections. The recent NSF Postdoctoral Research Fellowship on Interdisciplinary Research Using Biological Collections could be extended and expanded to include other aspects of biological collections management, care, and use including curation, digitization, data management, and education.

While many biological collections are used in undergraduate and graduate education, there is enormous potential to bringing more students and postdoctoral fellows into the collections for both education and research opportunities (Kreuzer and Dreesmann, 2016). Exposing the public to biological collections involves a substantial amount of additional work. However, behind-the-scenes collections tours were increasing in popularity (until the recent pandemic) and can provide not only educational opportunities for the public but also potential fundraising schemes. There are many examples of these types of events, such as the behind-the-scenes tours of the Field Museum in Chicago, Illinois (Golembiewski, 2016), and the Natural History Museum of Utah,[14] and many biological collections could make more effective use of their vast collections. One potential opportunity for supporting graduate students in biological collections is the NSF GK–12 Graduate STEM Fellows in K–12 Education, where graduate student researchers are supported to interact with K–12 educators.

CONCLUSIONS

Cultivating a highly skilled collections workforce, one that serves the data-intensive, globally connected, and often fast-paced needs of science and society, is essential to the long-term sustainability of the nation's biological collections.

[12] See https://mvz.berkeley.edu/undergraduate-program.
[13] See http://museumstudies.ku.edu/internship.
[14] See https://nhmu.utah.edu/events/behind-scenes.

The collections workforce is as important to a biological collection as the physical infrastructure and cyberinfrastructure. If biological collections are to not just survive but thrive throughout the 21st century, they will need effective, visionary, and well-supported leaders in addition to competent and innovative scientists and educators. Therefore, the workforce pipeline cannot be an afterthought; it requires consistent attention, planning, resources, and ongoing, dedicated stewardship. Truly, the question is not whether the biological collections workforce requires intensive investment, but how best to provide it.

There are still many unknowns about the biological collections workforce—its size, scope, diversity, and impact on the scientific enterprise. Careful assessment on a periodic basis would help fuel comprehensive thinking about current and future workforce needs, particularly the structure and function of a workforce pipeline that enables students to prepare for and connect to biological collection careers and supports training and professional development of existing biological collections experts.

The challenges facing biological collections are beyond the capability of any one institution to adequately address. A deeper understanding of the scope and needs of the existing collections workforce, identifying critical skillsets shared among the nation's biological collections, and building a sufficient workforce pipeline requires collaborative, coordinated action. The path forward will require collaboration among the nation's biological collections as well as partnerships with other professional communities, incentivized by the support of NSF.

RECOMMENDATIONS FOR THE NEXT STEPS

Recommendation 6-1: The leadership of individual collections, host institutions, relevant professional societies, and collections funders should collaborate to develop and strengthen the workforce pipeline through community-level action on the following issues:

- *Critical Skills.* Define critical, broadly applicable skillsets needed to lead, manage, and care for biological collections and expand and promote their uses for the national and global scientific enterprise and the benefit of society.
- *Workforce Analysis.* Conduct a comprehensive analysis of the existing collections workforce that, at a minimum, examines the professional responsibilities, demographics, education and training, incentives, compensation and benefits, and perceptions of greatest needs and opportunities for career development. Such an analysis should be conducted on a periodic basis (e.g., every 5 to 7 years) to inform community-level conversations and strategic action plans.
- *Diversity, Equity, and Inclusion.* Develop and implement programs to build a more diverse, equitable, and inclusive workforce. These programs should include elements such as restructured classroom and mentoring practices, student internships, research opportunities to ensure that opportunities are more visible and accessible to diverse students

and early-career professionals, and dedicated funding programs for internships and conference travel, workshops, and mentoring programs for diverse students and early-career professionals.

- *Education and Training Coherence*. Harmonize the design and offerings of biological collections–focused curricula and certificate and degree programs to fill current and future workforce education and training needs. This effort should include developing partnerships and cooperative arrangements with professional societies (e.g., for collections management training and taxonomic expertise), professional networks (e.g., in formal and informal education), and professional programs (e.g., museum studies, library studies, data science), respectively, to facilitate the design and implementation of biological collections–focused education and training programs in skillset areas not traditionally part of scientific training, and creating an online registry or portal to facilitate centralized access to information sharing about available education and professional development opportunities.
- *Alternative Staffing Models*. Provide guidance on alternative, innovative staffing strategies, including mechanisms to formalize student or volunteer involvement in collections management, that can help address staffing shortages, meet critical skillset needs, and serve as a mechanism to deepen collections knowledge among a broader range of people.

Recommendation 6-2: As part of its programmatic endeavors to promote a robust biological infrastructure, the National Science Foundation Directorate for Biological Sciences should support initiatives that focus explicitly on systemic, systematic, and thoughtful development of the biological collections workforce pipeline. In partnership with other directorates, such a programmatic focus should encompass future (e.g., students and postdocs) and existing collections personnel (e.g., early-career and senior curators and collection managers), and be predicated on maintenance and growth of biological collections infrastructure to meet diverse needs of societal import.

7

Securing Financial Sustainability

Long-term financial viability is critical to the ongoing and growing use of biological collections for research and innovation. Maintenance and replacement of aging physical infrastructure, continual upgrades to cyberinfrastructure, additional personnel to manage growing digital resources, upgrades to meet the needs of new emerging types of collections, new quality standards, and evolving requirements for permits and safety regulations are some of the funding needs that, while essential, may go beyond what annual budgets have covered historically. Sustainable resources for normal operations and upgrade costs can be found if collection and institution leaders can leverage support from a wide funding base. Central to this effort is communicating the role of collections and placing them as critical infrastructure that can benefit society. Collections need an adequate, predictable flow of resources to maintain the specimens and the data that are their historical legacy, while also innovating and adapting to new uses and demands.

Achieving financial sustainability is a goal for all institutions with biological collections. Financial stability is not just funding daily operation; a collection on a firm financial footing also has a source of funds for periodic building and cyberinfrastructure upgrades, new technologies that enhance the management and sharing of data, fieldwork, and salary adjustments and professional development opportunities for staff. At a minimum, collections require a sufficient annual budget for staff and supplies so they can follow best practices for storage, curation, growth, and access to collections, and can fulfill user requests for data, physical loans, and acquisition or in-person visits. Living and natural history collections with a secure financial future are able to focus their efforts on finding new ways to leverage their holdings for research and education, as well as supporting the addition of new specimens. These may include those collected in traditional ways and "next-generation" collections (Schindel and Cook, 2018) that may cross taxonomic and preservation-type boundaries. Such collections are also able to extend their specimens through linkage to derived products such as gene sequences and tissue collections (Hazbón et al., 2018; McCluskey, 2017; Rabeler et al., 2019). Overall, collections with sufficient resources are best suited to support basic biodiversity research and pressing societal challenges such as food security, climate change, invasive species, infectious diseases, and agricultural productivity in a rapidly changing global ecosystem. Institutions need to identify new strategies for sustaining and growing collections along with access to their data. New users must be engaged while anticipating, adapting to, and taking advantage of new funding models and sources to respond to changing needs and pressures. Reaching new partners and audiences requires developing new communications and networking strategies geared toward placing the collections at the center of all projects and activities. This will strengthen existing connections while building new ones with a diverse range of educational, scientific, corporate, civic, nonprofit, and government organizations.

CHALLENGES

Collections without funding and strategic planning to support physical and cyberinfrastructure, quality control, and personnel infrastructure will inevitably lose their ability to engage students, users, and members of the public via educational opportunities, or make contributions to the common good via transformative research. As the needs for research, education, and the expanding end-user community for biological collections increase, so does the pressure for long-term financial stability. For many public academic institutions, federal, state, and county financial support can represent a large part of the collections support, which makes them vulnerable to fluctuations in public funding. In addition, there are few long-term funding models for infrastructure. Because collections vary in their sizes, types, and missions, it is often difficult to apply a successful funding model from one institution to another.

Evidenced by the number of recent collection closures and troublingly high collection-to-support staff ratios (Thiers et al., 2018), today's biological collections are not stimulating the funding needed to sustain a vibrant and innovative collective resource that meets the needs of its user communities. So, is there a "business model" that can sustain the long-term viability of living and natural history collections under this common constraint? Here the committee describes some of the most typical and pressing challenges.

Short-Term Versus Long-Term Support

Biological collections are a long-term distributed infrastructure in support of research and education. As well as maintaining operations in the near term, collections and their institutions need to be able to anticipate future trends and changes in methods, technology, research applications, and regulations that may affect the maintenance and long-term sustainability of collections. While federal and other agencies have provided millions of dollars to fund short-term research projects that generate or use collections, the difficulty to assess the national portfolio of biological collections and the lack of a complete catalog of specific collections and specimen holdings make it difficult for funders to determine whether and how to spread their support. Biological collections are also lacking a clearly outlined long-term mission that is easily understood and inspiring. Because it is a distributed network of individually funded collections, it is hard to get the general public to support biological collections by demonstrating their role in describing and understanding life on Earth as well as patterns of diversity and extinction. On the other hand, the National Aeronautics and Space Administration is successful in obtaining the general public's support because it clearly describes one of their major goals: to land a human on Mars. Driven by measurable impact, agencies often face a tension between funding cutting-edge research where the benefits are easy to envision and quantify in the short term versus physical and digital infrastructure where benefits may be less obvious, less tangible, and long term. Thus, despite occasional federal support for improving infrastructure—and recent temporary funding to support the digitization of natural history collections—these collections need to generally rely on institutional funds for ongoing operations. In addition, living collections rarely have local institutional support, and a long-term federal strategy to support our nation's biological collections has not been developed. A clear, long-term vision for both individual collections and the collections community is needed for successful fundraising. As many collections continue to struggle to meet short-term basic needs of curation and infrastructure support, long-term financial stability is needed to ensure continued access to high-quality specimens and data and ongoing innovation in curation and data use. Decades of effort by both collections professionals and the extended research community could be lost if funding for a collection is put on hiatus or discontinued.

Limited Funding and Limited Pool of Funders

Biological collections require perpetual financial support to fulfill their mission for research and education. Collections are expensive to build and operate, as is retaining highly skilled collections staff. For most collections, these ongoing maintenance costs need to be funded from annual operating budgets provided by their institution. Such budgets, especially in the case of not-for-profit institutions, may barely cover the ongoing costs and are often subject to cuts or reallocations to other activities. Even with consistent annual operating budgets, collections will have unmet financial needs when faced with needs for upgraded and expanded cyberinfrastructure, new health and safety regulations, and unfunded mandates such as legal and regulatory compliance. For example, it is important to be able to respond to the implementation of new legal requirements that may affect collection growth and existing protocols such as the Convention on Biological Diversity, which includes the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization (see Box 1-10), the International Treaty on Plant Genetic Resources for Food and Agriculture, the Paleontological Resources Preservation Act, and others. A collection needs to be able to adapt to an increasing financial burden and legal operating requirements of such new collecting, acquisition, and dissemination practices.

When institutional funding is insufficient, collections seek external funding support to improve and expand collections, and sometimes even to fund basic collections care and infrastructure. For example, the National Science Foundation's (NSF's) Collections in Support of Biological Research (CSBR) program specifically funds biological collections infrastructure. The National Institutes of Health (NIH), through its Office of Research Infrastructure Programs, funds diverse living stock collections that support health research, which include vertebrate and invertebrate organisms. Research grants typically do not include support for collections beyond the processing of the collections made for a particular research project. The most cost-effective time to curate and digitize a collection is when it is first obtained. Costs for corrective measures or to deal with backlogs increase with time, which often is not taken into consideration or not funded.

A collection that is insufficiently funded to maintain its infrastructure will fall into a downward spiral in which use, ability to accept new material, quality control, and curation best practices all diminish, further limiting the institution's ability to obtain funding. Collections must face the challenge of how to communicate their mission to a diversified pool of funders such as public funders, institution leaders, and private donors in order to obtain sufficient infrastructure support.

Underappreciation of the Value of Biological Collections

Although specimens from biological collections are being used in a broad range of educational endeavors (see Chapter 3) and modern research such as studies of climate change, species interactions, and functional traits, as described in detail in Chapter 2, as a community, biological research collections do not market themselves well or effectively demonstrate their value to stakeholders. The centrality of biological collections to these educational and research activities is still not widely appreciated outside of the immediate research and collections community, as evidenced by the recent defunding of active collections at places such as The University of Oklahoma (Nhcoll-l Listserv posting from Dr. Dan C. Swan, Interim Director, Sam Noble Museum of Natural History) and the University of Alaska (Lambert, 2019a). More broadly, the fact that many collections in the country are understaffed (evidence from herbaria in Thiers et al., 2018) is consistent with poor messaging about the successes of this infrastructure. For living collections, this underappreciation of the potential applications of the collections leads to the lack of financial support for collections used for research, especially for smaller institutions, which leads to increased users' fees, a decrease of collections use (Kevin McCluskey, Kansas State University, personal communication, 2019), and loss of competitiveness. Many researchers who make collections for their research across the breadth of the biological sciences in fields such as ecology are sympathetic with the goals of collections, yet for logistical, lack of awareness of the collections infrastructure, or financial reasons may not contribute to the deposition or accessioning of their specimens in collections. This failure to deposit specimens or samples made into the appropriate collections results in "dark data" (see Chapter 5) and will severely limit the impact of collections in the future. Additional funding to ease the significant burden and cost of specimen deposition and accessioning, as well as a change in culture within the biological sciences, will help ameliorate this gap. There are some exceptions (George, 2019), but generally, collections fall back on traditional value propositions (Merritt, 2017), which are not necessarily compelling to modern funding sources.

Communicating Outcomes and Impacts

Successful metrics for outcomes and impacts can be critical to continuing support, both at the institutional and the community level. The challenge that arises is agreeing on what metrics are important to share with stakeholders within and outside the collections community. As with many scientific endeavors, collections management lacks a common set of metrics that can be aggregated across collections, especially given the great variability in the collections landscape in scale, scope, and material. This is made more difficult by the diverse and ever-growing body of stakeholders making decisions about collections, who have dynamically shifting priorities and requirements. There is a need to share observations and conclusions in ways that people can understand and through multiple channels of communication. The benefits of more effective communication include:

- Enable biological collections administrators and staff to articulate the vitality of their programs.
- Demonstrate use, impact, or other dimensions about how a collection is aligned for the needs of the host institution, science, and education communities, or funders.
- Enable the host and funding institutions to learn about and assess the returns on their investment.

Collections typically collect metrics that they believe their funders or host institutions want; therefore, the metrics used by one institution for its own funders and host institution might not be useful for other institutions. If these metrics are used for all institutions, this might mask the value of the contribution of other collections. This may render the evaluations unhelpful or even disadvantage individual collections when metrics are aggregated across different kinds or locations of collections. The community needs to find a set of metrics that can be used to demonstrate the contributions of a broad range of collections of different sizes, types, with different objectives. Individual institutions need to also be able to articulate how they "rate" relative to other, similar collections or collection endeavors when trying to communicate or compare impact to funders or organizations when faced with competing requests for support. A comprehensive set of metrics would help to assess, compare, and then communicate impacts from different kinds of approaches to collections more effectively.

Estimating the Financial Value and the Cost of Biological Collections

Summaries of the many ways in which collections are valuable for research and education have been described several times (Allmon, 1994; Anderson, 2012; IWGSC, 2009; Meineke et al., 2018a; Nudds and Pettitt, 1997; Suarez and Tsutsui, 2004) and also in Chapters 2 and 3. However, better metrics are still needed to assess the importance of collections in monetary terms at a time when economic value alone often dominates our public discourse. Estimates of the cost of specimens, and therefore financial value, from field collections to data entry have been conducted for individual collections. For example, Bradley (2012), as well as Baker et al. (2014), assessed the cost for collecting and housing their mammal collections at an average of ~$70 per specimen. For living stock collections, the cost of a specimen is reflected in users' fees. For example, acquiring a specimen from the Fungal Genetics Stock Center collection costs between $25 and $50[1] but would reach 300 euros when the cost of accession is taken into account (Smith et al., 2014). However, these costs do not include up-front capital cost, facility maintenance, and all other activities pertaining to services to the research and education communities. While the cost of conducting several years of field studies leading to the collection is high, it is generally acknowledged that maintaining these collections annually costs a relatively smaller fraction of that amount. Smith et al. (2014) outline budget models for microbial biological resource centers, including the estimated cost to preserve, maintain, and distribute each specimen. Considering the extraordinary range of types and purposes of biological collection specimens, these knowledge gaps make it difficult to articulate for funders or administrators who are not scientists the value proposition of biological collections in ways that translate into increased resources. Describing assets and articulating the return on investment for collections is also difficult to calculate, and the financial consequence of the unavailability of specimens or their associated data is a question that is at best challenging to answer or one that is rarely even asked. This needs to be recognized as the opportunity cost of not having a well-preserved collection when information contained in such a resource is needed (Freedman et al., 2015).

Financial Obligations for Ongoing Growth of Biological Collections and Their Associated Digitized Data

In addition to the ongoing growth of collections in size and numbers of specimens, the advent of new technologies and associated research endeavors increase their accessions and diversity of uses (Wandeler et al., 2007). Collections are often housed within not-for-profit institutions such as colleges or universities, which need to balance many commitments in terms of space and personnel. Support by these institutions for any obligation, including acquiring and maintaining collections, will vary over time and is contingent on the obligation aligning with the mission of the institution. Ideally, institutions commit, either implicitly or explicitly, to a business strategy that commits to the growth and maintenance of collections in their care. However, with limited funding and staff members, hard conversations about deciding when a collection needs to be closed, discarded, or transferred to free space for other activities are vital. Instead of discarding a collection that no longer has a clear relationship to the evolving mission of the institution, or is no longer deemed important enough to maintain locally, near to other university activities, it is not uncommon for

[1] See http://www.fgsc.net/fgsc/pricing.html.

colleges and universities to cease supporting or maintaining the collection while still retaining it on-site. These "orphan collections" eventually may become damaged from inattention beyond the ability to save them. Sometimes these collections are offered to another institution that can absorb it into their holdings. However, the collections community is made aware haphazardly that a given collection is on the brink of being lost, such as what occurred with the fish and herbarium collections of the University of Louisiana Monroe (Flaherty, 2017). In such cases, the community can rally behind saving the collection and identify a suitable repository that can absorb the collection. One of the criteria for funding by the CSBR program at NSF is urgency, such as when a collection is in danger of being destroyed or lost due to a failure of the host institution to continue its financial obligations to sustain the collection. Typically, only larger institutions and collections are able to absorb collections that are in critical danger of being lost, a process that can lead to unusually rich aggregations of collections that may better serve specific research communities. Most collections, however, typically lack the funds to accession large numbers of new specimens. Rescuing biological collections is particularly difficult for collections that do not have explicit local support. The living collection community can rescue endangered or orphaned collections, but only when there is sufficient existing capacity (Boundy-Mills et al., 2019). In other cases, taking on orphan collections puts undue pressure on existing infrastructure and funding. The collections community and funders need a strategic vision that includes a variety of tactics and benchmarks to prioritize accessions and deaccessions of collections and to alert the community of collections in danger of being lost.

Financial commitments and strategic planning to continue to digitize specimen records and build and maintain the cyberinfrastructure are also required to ensure the long-term utility and accessibility of digital data associated with biological collections. Digitization and a strong cyberinfrastructure provide online access to specimen-related resources and increase opportunities in research (see Chapter 5). However, digitization is a time-consuming endeavor that necessitates trained staff members to manage at least these major tasks: (1) digitize already collected specimens; (2) digitize new specimens that continue to be accessioned; and (3) regularly reassess the digital data needs of the user communities, including software and hardware needs for preserving, interpreting, and disseminating digital resources. Thus, it is especially worrisome that there is no long-term nationwide strategy to simultaneously support the high cost for generating digitized data and storage infrastructure for newly collected specimens, while at the same time retroactively capturing data gathered over past decades and centuries. These dual efforts for digitization will require new investments and planning for long-term support.

Lack of Business Management Training

There are only a few avenues for business management and fundraising training opportunities, and these are often limited to institutional leaders who frequently have short-term appointments. This pattern limits the ability of any given institution with rotating leadership to build and leverage new resources beyond traditional renewable funding. In addition, the staff is typically hired for research knowledge and curatorial skill, not for their business or financial acumen or administrative leadership. These issues are not unique to biological collections and affect any community with significant research, education, and infrastructure requirements beyond laboratories. Fee-based training programs,[2] which are often beyond the financial reach of small collections, and workshops are starting to address some of these issues (Parsons and Duke, 2013) but do not go far enough to address the ongoing and changing financial training needs of biological collections administrators. Collections scientists are routinely expected to become fundraisers, but this may be detrimental to other activities. Insufficient training to develop business models and financial strategies may lead to a "nonprofit starvation cycle" where institutional leaders may have unrealistic expectations about how much it costs to run a collection, which results in either not asking for what is truly needed to prevent losing out on receiving funds or by cutting corners on vital needs. As a consequence, funders have misperceptions about what collections truly need.

RANGE OF OPTIONS FOR ADDRESSING THE ISSUE OF FINANCIAL SUSTAINABILITY

Overcoming the barriers described in the previous section will require a thoughtful approach that takes advantage of resources that are available not only within the biological collections community but also from outside the community. This section describes a number of strategies for surmounting these barriers, including developing strategic business models and long-term frameworks to diversify funding portfolios and explore diverse funding mechanisms, strengthening partnerships and offering training opportunities, taking advantage of well-established communications practices from the science communications community, and developing a national vision for ensuring financial sustainability.

Developing Long-Term Strategic Frameworks for Building a Diversified Funding Portfolio

Most collections are utilized for research and education and obtain their funds from single sources. Strategic planning helps identify the financial and other needs of a collection, suggests areas of potential savings, and differentiates the funding needed for ongoing maintenance of the collection from what is needed to meet evolving standards, replace aging infrastructure, and accommodate the growth of collections. Going through the strategic planning process every few years can help identify the potential funding sources for biological collections infrastructure and also identify gaps in funding that will need to be met by other resources during the plan's duration (Parsons and Duke, 2013).

---

[2] See https://www.esa.org/programs/training.

For example, NSF's CSBR and several funding programs through the Institute of Museum and Library Services offer grants that may offset the costs of the improvements needed to maintain adequate infrastructure for collections survival. Research and education initiatives using collections may be funded by a wider range of public and private sources than are available to support collections infrastructure, and thus it is imperative that all such initiatives cover the full cost of that use.

However, the need for major continued infrastructure improvements at all U.S. biological collections is not being met through grant programs alone. Developing a diversified funding portfolio (and subsequent fundraising) is a desirable outcome of a strategic plan, one that contains a mixture of institutional operating funds and funds that are raised specifically for the collection, for example, endowments; user fees for partial recovery of service expenses (where allowable and practical); licensing of images for commercial use; donations from alumni, members', or friends organizations; project-based grants and contracts; and sustaining grants including naming opportunities from philanthropic or commercial enterprises. A diversified funding portfolio built on stable base funding will help ensure a collection's sustained security and viability. The Natural History Museum of Utah (NHMU) is an example of an institution with collections that garnered corporate support for a state-of-the-art facility. During a presentation to the committee in 2019, Dr. Sarah George, then executive director of NHMU, outlined a strategic framework that included providing training to collections staff members in building long-term fundraising strategic plans, developing case statements with stories that appeal to donors, and establishing a community of practice led by collections professionals with experience in fundraising. Similarly, the William & Lynda Steere Herbarium of the New York Botanical Garden is an example of a collection that has been endowed through individual philanthropy. The named Bayer Center at the Missouri Botanical Garden reflects the corporate support obtained for that collection building in the late 1990s. Various financial models have been explored for the maintenance of biological collections and data infrastructure (e.g., Chandras et al., 2009). A strong and stable base requires recognition of the value of the collection to the mission of the larger institution. Above all, a collection's leadership needs to ensure its appreciation as critical infrastructure that supports the institution's research, educational, and other goals.

Building Funding Partnerships

Given the critical role of collections across a range of scientific disciplines, funders need to take advantage of opportunities to tie funding for collections infrastructure and cyberinfrastructure to other infrastructure investments and initiatives across agencies. An example is a collaboration among the world's major herbaria, the Mellon Foundation, and JSTOR to provide Global Plants, a database of approximately 2 million type specimens of plants and fungi. Between about 2004 and 2015, the Mellon Foundation funded the digitization of type specimens from about 200 herbaria worldwide. Images were the property of the institutions holding the specimens imaged, but a copy of each image and its associated metadata were added to the Global Plants database managed by JSTOR and offered as part of a subscription package to libraries and herbaria worldwide. The subscription fees support the maintenance of the database and the contextual linkage of these type specimen records to other JSTOR holdings (JSTOR Global Plants[3]). Other examples include collaborative networks to develop and support software initiatives such as Specify and Arctos, and Thematic Collections Networks funded for collaborative digitization projects through NSF's Advancing Digitization of Biodiversity Collections (ADBC) program. Another example involves ownership of specimens by one body (e.g., the Bureau of Land Management), curation by a university-based or stand-alone collection, and infrastructure support by another body of funding. Under this model, funds appropriate to the number of specimens would need to be provided by the appropriate agency to the institution housing and curating the specimens. Partnerships between federal agencies and non-federal sources, such as foundations, could also be explored as possible resources for supporting collections as infrastructure.

Communicating: Working on the Messaging

Collections are constantly being accessed, curated, annotated, measured, photographed, used for research, and cited, and each specimen added to a set subtly expands the scientific and educational uses for which the collection can be engaged. Establishing and communicating the relevance of biological collections will ensure that they are considered as an essential element of the fabric of the institution. Biological collections are most appropriately envisioned as research centers, many of which have public displays for formal and informal education. Universities spend millions of dollars on research centers, such as building new spaces that allow professors and students to leave their department silos and engage in interdisciplinary, transdisciplinary work (see Chapter 4). Biological collections can also bring communities together and their value needs to be communicated as such. For example, if a collection is part of a university, the use of the collection by multiple departments—biological and beyond—and other units will help ensure relevance. Moreover, engaging with students, alumni/ae, and others who use the collection can strengthen its position within the institution. Establishing and communicating the role of the collection in the local community will enable it to build community interactions, such as developing a strong volunteer base, providing opportunities for citizen science, and other initiatives (George, 2019). Working with development officers to raise funds to establish and grow an endowment is crucial. Collections need to seek to benefit from larger institutional capital campaigns.

These efforts to build relationships with various communities require that the collection introduce products that address the emerging needs of the relevant stakeholders and track activities to show impact. Biological collections serve many needs and many stakeholders, each of which needs messages and

[3] See https://plants.jstor.org.

narratives that resonate specifically with them. For some of these stakeholders, robust metrics and data may be persuasive or compelling. But many of the defining benefits of biological collections, such as serendipitous uses and new discoveries, are best documented as descriptive narratives about advances in knowledge and other types of success (IWGSC, 2009). These narratives include research, educational, and public service contributions, some resulting in outcomes that can only arise through the use of biological collections (see Boxes 2-1, 2-2, 3-1, 3-2, 4-1, 4-5, and 5-1). A well-developed literature and an established community of practice for the science of science communication (Jamieson et al., 2017a; NASEM, 2017a) support the development of compelling narratives—often best told retrospectively—that identify areas or problems that were solved or elucidated by access to biological collections and their associated datasets.

The biological collections community does not need to reinvent the wheel to find ways to develop, structure, and describe its successes. However, developing a set of guidelines within the collections community for how to develop clear narratives, what topics are best suited for narratives about success, sharing experiences with how and when to tell these stories, and compiling a community-wide list of these contributions can synergize with formal evaluations of the value and impact of biological collections.

Demonstrating Return on Investment and Benefit–Cost Ratio of Biological Collections

Investing in scientific research and education pays off. This is a valuable component of sponsor stewardship regardless of the kind of reporting required. Demonstrating the return on investment of collections in support of research and education is more difficult and somewhat anecdotal. It requires some careful analysis complementary to, but different from, demonstrating the impact of collections on research and education. For both, the metrics involved would, therefore, be different (see Chapters 2 and 3). According to Dr. Keith Crane from the Science and Technology Policy Institute, the best-documented examples of benefit–cost analysis come from the agricultural communities that can estimate in dollars crop production and productivity after an intervention using biological collections (Crane, 2019). The cost of financial consequence for not sampling the environment for emerging and re-emerging pathogens can also now be estimated. Recently, reports described that the World Bank mobilized more than $1.6 billion for Ebola recovery and estimated that the region's gross domestic product would lose $2.2–$7.4 billion over the short term. This story could have been different if relatively small funds were made available for collecting field samples and identifying the local distribution of the viruses after the initial discovery (DiEuliis et al., 2016). According to Merritt (2017), rethinking the value proposition of biological research collections will be key to ensuring their financial sustainability. Finding ways to increase appreciation for the invaluable contributions that biological collections make to research and innovation will be the first step in ensuring their health and stability in the future.

Developing Strategic Business Models

An individual collection can do a lot to improve its financial sustainability. Foremost is developing a comprehensive annual budget to ensure optimal operation, guided by an up-to-date strategic plan that is periodically reviewed and updated. Financial management of a collection needs to incorporate business models, develop relationships with relevant stakeholders and funders, and, if part of a larger organization, connect to that institution's mission and goals. Specifically, a collection's leadership needs to include expertise in business, finance, marketing, and networking, as well as biology, Earth science, and data science, among others. The business model needs to account for infrastructure (acquiring, maintaining, upgrading), but also adapting to personnel capacity and needs, specimen loans, including use for on-site research, education, or others. Importantly, the business model needs to include some type of marketing or outreach programs in addition to strategies to grow collections either by acquiring or integrating new accessions. Finally, the plan would also include a comprehensive risk management plan for fire, natural disasters, shutdown, or infrastructure failure (e.g., burst pipes, failure of temperature or humidity control systems). This includes:

1. Articulating expected outputs given the objectives of the collection and the needs of the community. Outputs are more than research publications. They need to be tied to infrastructure as well. For example, is growth an expected output of the collection? Then, collecting and accessions need to be taken into consideration.
2. Determining the appropriate level of funding diversification and identifying all possible revenue streams. The level of diversification has to be aligned with the expected outputs of the collection.
3. Articulating the key sustainability elements of the collection. What absolutely must be in place and appropriately funded for the collection to be able to deliver on its objectives in a sustainable manner [without burning the candle at both ends]? For most research centers—a good parallel—sustainability elements are facilities and equipment; operational personnel; and research and researcher support personnel.
4. Determining what approach needs to be taken on core funding—the pool of money that a collections director [or board of directors] can allocate where it is needed for operation or even for exploring new ideas/capabilities for the collection.

Complementary Funding for Research and Infrastructure

Only a few funding sources—most notably NSF and NIH—invest in collections infrastructure, although support may also come from other federal agencies, state and local agencies, foundations, collection-holding institutions, and individuals. Ideally, support for collections infrastructure needs to be seen as an underlying requirement of the research being conducted and not as coming at the expense of support for research. For example, NSF funding to support the accessioning and digitization of specimens collected as part of an ecological or evolutionary study would generate funds for collections to perform the tasks necessary to make these specimens available while providing a foundation for innovations in research and education. An NSF-funded and mandated specimen management plan (see Chapters 4 and 5) would provide the necessary guidance and structure to require housing of specimens in appropriate collection repositories for specimen-based research. This plan would promote communication between researchers and the collections where the specimens and their associated material would be deposited. Because this would happen during proposal preparation and before collecting, it would allow the exchange of information on data collection, best practices, and protocols to maximize specimen and data quality and help identify taxonomic and geographical gaps among others. The plan would help link research funding to collections by mandating per-specimen funds in all specimen-based collecting proposals necessary to curate, digitize, and provide long-term care of those specimens. The collections community could provide guidance on such a specimen management plan for adoption by NSF.

Training and Sharing Best Practices

Collection professionals often lack expertise in business models and financial planning and training on topics such as developing an accurate budget or exploring innovative ways in diversifying revenue streams. One-off financial sustainability workshops convened by the American Alliance of Museums (Merritt, 2017) focusing on natural history collections and the Ecological Society of America (ESA) focusing on living collections (Parsons and Duke, 2013) have demonstrated the need for more training in this area. ESA has started offering annual training programs that focus on funding (see ESA's Sustaining Biological Infrastructure [SBI] Training Initiative[4]). While there has been little overlap between natural history and living collections in this arena, both communities have taken different approaches to financial sustainability but have insights to offer each other. Developing a network of museum directors and biological collections administrators across collection types, who can share best practices for financial models and planning, would have a more extensive impact on the biological collections community as a whole (see Chapter 8). Networking between collections' directors and representatives of funding institutions is an opportunity to increase the limited funds currently available for supporting collections and to develop novel funding mechanisms.

Willingness to Pay

Specimens held in natural history collections and their digital information, both of which are often irreplaceable contributors to educating generations of

[4] See https://www.esa.org/programs/training.

scientists and advancing basic and applied research, have a history of being available to most users free of charge or at a minimal cost. Part of the explanation for this arrangement is that not-for-profit institutions hold most biological collections. Also, specimens are often collected with public funds through NSF, and therefore these institutions cannot always justify charging for their use. Traditionally, natural history collections exchange, borrow, or lend specimens within their communities on a quid pro quo basis. But for just about any other "service" in the world, people pay for that privilege. Innovative solutions may require bringing in social science research to assess how user fees do or do not fit into business plans for collections. Ultimately, there needs to be a fundamentally different funding paradigm for collections to be maintained and thrive. Lessons learned from the advent of paywalls in the print media and the creation of journal consortia could afford examples for biological collections working to adapt to a rapidly changing funding landscape while working to establish new models for support and partnerships.

The business model is different for some living stock collections, which have used subscription and fee-for-service plans for decades. Lower fees are usually applied for educational and research use. It is worth noting that The Arabidopsis Information Resource was once a public resource and is now pay-for-access (Reiser et al., 2016). The American Type Culture Collection (ATCC) is also a nonprofit institution that became self-sufficient by charging for its materials and services (NRC, 2011). Both of these examples could provide insight into the pros and cons of a transition to a subscription-based funding model. Some collections do not loan or distribute their specimens and material to for-profit users, but for those that do, fees are traditionally only charged to (or will be higher for) for-profit users who will be making a profit based on the data extracted from the specimens or their metadata. For this reason, many living stock collections protect their rights and the rights of the donors of their material. Material transfer agreements (MTAs), or similar types of agreements (limited use licenses, for example) serve these purposes. The MTA limits the users' ability to transfer the material to ensure the quality of the collection's materials and encourages the primary user or other subsequent users to procure the material from the collection for a fee. Resulting funds may be used to further research or to sustain the ongoing operations of the collection. Typically, the MTA restricts usage of the material to research use only, and some collections may require a license to use the material for clinical or commercial purposes. Often, this model relies on the willingness of the users to approach the collection or the depositor to request a license. It also depends on a collection's capability to track the use of its material, which often is a challenge and can sometimes only be accomplished by larger collections, such as ATCC, with a license department that can monitor the clinical or commercial use of their material. Natural history collections have made much of their data available free of charge online in the past decade. Undoubtedly many for-profit companies have taken advantage of those data, but currently this landscape is unknown (see Chapter 5).

In the event a collection decides to change its funding strategy, it will be essential to bring in the expertise required to learn how to charge for its materials/

usage, improve accessibility to its users and improve its websites and customer support to establish a solid platform to make the collection profitable. This is not an easy transition and can take time to develop and establish and might require an important investment of funds. It is here where funding entities could provide the necessary investment to implement these changes.

Develop a National Vision for Biological Collections and a Distributed Collection Network in Service to the Nation

The collections community needs to assume a leadership role in developing a national vision for ensuring the financial sustainability of biological collections. While institutions that curate, maintain, and use biological collections have different missions, sizes, and purposes, they all face a complex balancing act to adapt to the evolving needs of biological collections. Unity within the collections community is fundamental to solving these challenges (see Chapter 8). Multidisciplinary research has blossomed over the past several decades. Positioning biological collections and their associated metadata as a key resource for addressing societal problems, such as the loss of biodiversity, global change, emerging infectious diseases, antibiotic resistance, and food security, would appeal to the many funding agencies with visionary research agendas.

Working in partnership with other collections could be a successful strategy to raise funds, reduce costs, or pool resources, especially for small collections. Networks and consortia of collections have been very successful in obtaining funds for digitization through NSF's CSBR and ADBC programs; local foundations or governments could be amenable to supporting ongoing or project needs of regional collections if they were confederated in some way, perhaps to support a local or regional biodiversity initiative. Collections consortia could also possibly reduce costs through shared supply orders to reduce unit costs, or by sharing equipment or other infrastructure (Parsons and Duke, 2013).

The community of collections professionals—professional organizations, staff, and faculty members at institutions of all types and sizes—is a powerful resource that can provide guidance, training, and support across a range of issues (see Chapter 6). Through various means, the collections community reaches out to help struggling collections through letter-writing campaigns to collections institution administrators (e.g., to administrators at the University of Alaska in 2019), or temporary adoption of imperiled collections. In 2015, the New York Botanical Garden's William & Lynda Steere Herbarium made room for the herbarium from the Brooklyn Botanic Garden, whose building infrastructure was in a critical state of disrepair. The collections are protected and made available for study in their temporary home until they are either returned to their original home or ownership is formally transferred to the Steere Herbarium. Networks of collection professionals can play an important role in catalyzing the development of community-wide initiatives to benefit the wider collections community.

CONCLUSIONS

The importance of the nation's biological collections to research and education calls for robust mechanisms to ensure their long-term financial stability. Physical infrastructure, cyberinfrastructure, workforce, and the evolving requirements for quality, accessibility, and usability of specimens and their associated data place growing financial demands on biological collections. The capability to not only maintain this infrastructure, but upgrade it to meet the multifaceted needs of science and society hinges on adequate funding. Central to this effort is the development of comprehensive business plans that include estimates of the public funds needed to support the research that generated the collection and the infrastructure needs of the scientists that use collections as well as maintaining and providing access to the collections.

Yet, not all biological collection leaders have sufficient expertise or support to develop comprehensive funding models, cultivate donor relationships, and engage the community of scientists and professionals who benefit from biological collections. Thus, efforts to identify new strategies for sustaining and growing biological collections will require both initiatives of individual biological collections as well as collaborative action of the biological collections community. A visionary collections community can accomplish this in two ways: develop compelling value propositions, business models, and strategic plans to implement and periodically assess their investments; and build partnerships to develop a national collections network to further the mission of collections in research and education. Researchers need to be encouraged to value not only the samples immediately relevant to their own research—and which may be lost to future researchers—but also the value of their specimens to future generations.

As documented throughout this report, biological collections produce a wide range of benefits for science and education in the United States and the global community. The financial sustainability of the infrastructure that provides those benefits, from individual biological collections to a network of collections to the full portfolio of the nation's biological collections, will require substantial attention, time, and expertise. Many individual biological collections do not currently have the resources to contribute to the comprehensive development of funding models. The biological collections community will need to act as one in order to develop partnerships, centralize a pooled set of data and resources, track the use of collections in research and education using diverse metrics (as described in Chapters 2 and 3) at the community level to show the national and international impact of U.S. collections, and identify new approaches to funding.

RECOMMENDATIONS FOR THE NEXT STEPS

Recommendation 7-1: The leadership (directors, curators, and managers) of biological collections should work with business strategists and communication experts to develop business models for financial sustainability and infrastructure of biological collections. Included in this discussion should be the development of a mechanism to:

- diversify funding portfolios and develop relationships with non-traditional partners who may provide collections support;
- assess a per-specimen acquisition and maintenance cost. This assessment would depend on the size and nature of the collection—both physical and digital; and
- explore revenue streams that could include pay-for-use models, the establishment of material transfer agreements and licensing systems, or perhaps pay for value-added for digital datasets configured for a particular purpose. Each of these approaches must be done in ways that avoid driving costs to levels that are prohibitive for researchers.

Recommendation 7-2: Professional societies should develop extensive networked training platforms for sharing best practices for financial management and planning and business models for collections of all sizes and types. This could be an ongoing activity centered at a national biological collections center and should include both natural history and living collections together.

Recommendation 7-3: The National Science Foundation Directorate for Biological Sciences should continue to provide stable, long-term funding to support investigators who rely on biological collections for research and education. Specifically, it should:

- work with other federal agencies to address research infrastructure support and needs;
- provide funding for the management and infrastructure of the collections themselves;
- collaborate with host institutions and other funders to establish new mechanisms and funding to collect, aggregate, and synthesize metrics to evaluate process and performance for biological collections; and
- support the accessioning, curation, digitization, and long-term care of specimens as well as the publishing of their associated data through a mandated specimen management plan.

8

Taking Collaborative Action

A sense of urgency informed the committee's deliberations, elevating the critical need to act on this report's recommendations now. Biological collections are vulnerable due to systemic underfunding and insufficient recognition of their importance to science, education, and society. This is at a time when the nation's biological collections are poised to be harnessed to provide data uniquely capable of informing challenges brought about by rapid and unpredictable global change. Unpredictable and unprecedented global changes have a huge impact on economies, health, and food security worldwide. The lack of knowledge of the identity, distribution, and interactions of biodiversity on our planet preclude our ability to predict or mitigate the emergence of pathogens (Cook et al., 2020; UNEP and ILRI, 2020) or understand the causes or consequences of the accelerated rate of species extinctions. However, collections are also vital for developing diagnostic kits, treatments, and vaccines. Pandemics and loss of biodiversity, however, are only a few of a growing number of threats to humanity due to changing environmental conditions that will urgently require more resilient and integrated initiatives to build and then leverage primary biodiversity infrastructure, such as the resources held in biological collections.

CRITICAL JUNCTURES INDICATE THAT THE TIME TO ACT IS NOW

A broad consensus of scientists has urgently emphasized that anthropogenic impacts, such as habitat conversion, overexploitation of resources, pollution, and climate change, are catastrophically challenging marine, freshwater, and terrestrial life (Ceballos et al., 2017; IPCC, 2019; Ripple et al., 2020). A growing and diverse set of alarming environmental metrics (e.g., increases in ocean heat content, ocean acidity, sea level, land burned in temperate and tropical zones, extreme weather, and decreases in the extent of sea ice, ice sheets, and glacier thickness) reflect extreme and cascading environmental changes now disrupting economies, public health, and the habitability of our planet. Understanding how these ever-accelerating changes will impact humanity has become a critical challenge facing the global scientific enterprise.

Biological collections stand alone in providing the temporal, spatial, and taxonomic sampling needed to document the effect of these changes on biodiversity in natural and managed ecosystems. Important clues to understanding, adapting to, and mitigating environmental changes reside in the living and natural history collections that are the focus of this report. Future efforts to manage and develop these biological collections need to be directed toward preserving existing resources for research and education. At the same time, new specimens must be added to fill in current knowledge gaps and new questions not even articulated. Designing rigorous programs that will allow us to understand, track, and mitigate impacts of changing global environmental conditions will require a renewed commitment to maintain and further develop the primary biodiversity infrastructure (i.e., specimens and informatics) held in biological collections. Future development of biological collections globally could more directly involve local communities and especially Indigenous populations, when possible (Colella et al., 2020; Cook et al., 2013), to promote engagement and reciprocity, including benefit-sharing, infrastructure, and capacity building. Natural history collections offer the ability to document and understand the rapidly changing biodiversity of our planet through time—in the present through new collecting, over the past few hundred years through existing collections (both large and small), and in deep time through fossil collections. Living stocks collections are important in this time of rapid change as, in addition to understanding changing environments (Ellison et al., 2011), they potentially hold answers to fighting new or re-emerging diseases (e.g., Zika, Ebola, coronavirus disease 2019) and developing crops that are more robust in the face of rapidly changing environmental conditions. To cite two of many possible examples, seed bank collections will be essential for identifying genetic resilience in crops, now largely monocultures, under disrupted climate regimes, and culture collections will be critical to characterizing emerging microbial pathogens and responding to threats to agriculture. With sufficient support, biological collections can offer not only a starting point for tracking and documenting change but predictions for the future use of modeling and artificial intelligence. As we enter a period of intensified research into documenting the response of ecosystems to change (exemplified, for example, by the National Science Foundation's [NSF's] Navigating the New Arctic program[1]), it will be more important than ever that biological collections continue to preserve specimens and share them and associated data on which scientific conclusions are based. Heightened awareness of the value that biological collections can add to virtually every facet of biology (and other scientific disciplines), and when coupled with sufficient resources to maintain and grow them, provides leverage to create the critical snapshot for this dynamic epoch. Collections provide the baseline infrastructure needed, not only for current and future research, but also to ensure environmental and societal resiliency.

Beyond changing environmental conditions, biological collections can also make transformative impacts on urgent societal issues by facilitating new collaborative ties among diverse disciplines (ranging from engineering to arts and humanities), ultimately stimulating new perspectives and creating synergistic initiatives. Dramatic changes in academic culture over the past decade favor integrative approaches to address complex questions. As detailed in Chapter 2, living and natural history collections serve a diverse array of research communities, which if brought together, hold great potential for interdisciplinary, broad, and synergistic endeavors to answer challenging new questions (e.g., global change, human health, food security) that necessitate teams of investigators pooling knowledge and working collaboratively. In an era of growing

[1] See https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505594.

interconnectedness, grand challenges of global importance, such as the United Nations Sustainable Development Goals,[2] call for a structured mechanism to bring people together as does one of the NSF Big Ideas, "Growing Convergence Research,"[3] which asks for a "deep integration across disciplines" and continues stating that "as experts from different disciplines pursue common research challenges, their knowledge, theories, methods, data, research communities, and languages become increasingly intermingled or integrated." To accomplish such integration, creative models for broad collaborations and networking among collections and institutions will be essential and need to be encouraged through funding cycles. For instance, to take advantage of the synergy of such collaborations will require a substantive realignment of federal financial resources, public infrastructure, and state and federal agency agendas through a better appreciation of how biological collections meet the mandates of federal public health and natural resource management agencies.

Biological collections are poised to make major contributions to today's burgeoning information economy. In addition to integrating across previously siloed disciplines from engineering to chemistry to biology, collections hold nearly limitless data, with each unique genome waiting to be explored, increasing our understanding of how they code for novel responses to environmental change and evolutionary adaptation (see Chapter 2). As described in Chapter 5, thanks largely to recent collaborative digitization projects that have helped build interinstitutional ties and opened up unprecedented access to the vast treasury of information they contain, collection institutions can now capitalize on their unique platforms (i.e., biodiversity sampling) to demonstrate how science can be integrated across disciplinary boundaries as collections continue to emerge as the central infrastructure for addressing a series of critical societal needs. More than ever, biological collections now have an energized community that is ready to step up to meet these grand societal challenges.

A Framework for Collaboration and Innovation Is Needed

It is clear that the time to act is now. This report, along with many others, details challenges facing living and natural history collections and what is at stake if biological collections collapse and collecting ceases. This report also offers an issue-specific range of options regarding physical infrastructure, cyberinfrastructure, personnel, evaluation, financial sustainability, and connecting to national priorities and needs for research and education. To ensure the long-term sustainability of individual collections, thereby strengthening the national portfolio of research infrastructure within the next decade, collaborative solutions to these challenges need to be developed and implemented. Throughout the report, a number of the committee's recommendations, however, require a unified vision and strategy—the biological collections community will need to embrace and implement *collaborative* action. E. O. Wilson (1998) made the

[2] See https://www.un.org/development/desa/disabilities/envision2030.html.
[3] See https://www.nsf.gov/news/special_reports/big_ideas/convergent.jsp.

intellectual case for this sort of thinking: "We are drowning in information, while starving for wisdom. The world will henceforth be run by synthesizers, people able to put together the right information at the right time, think critically about it, and make important choices wisely."

Several research communities have established central hubs, multi-tiered networks, associations, or synthesis centers, funded through NSF grants or other federal and state support, to explore innovative research and education opportunities through collaborative analysis and synthesis at facilities that provide computational and logistical support. The National Center for Ecological Analysis and Synthesis,[4] the National Socio-Environmental Synthesis Center,[5] the United States of America National Phenology Network,[6] the Association of Science and Technology Centers,[7] or the John Wesley Powell Center for Analysis and Synthesis[8] are a few such examples. Such centers and networks are considered critical research and education infrastructure, enabling the synthesis of data that cut across disciplines and perspectives to address societal challenges (Baron et al., 2017; Rodrigo et al., 2013). The biological collections community could leverage the organizational structure of centers and networks as a model to establish an Action Center for Biological Collections, whose mission would focus on all biological collections and offer a collaborative platform to provide actionable and lasting solutions for the collections community at large.

Although the biological collections community is motivated and active, many of the community's endeavors to communicate the role of collections and position them and their associated metadata as critical infrastructure for addressing societal problems are disconnected and uncoordinated. A collaborative action center would facilitate and connect all relevant and interested parties, including living and natural history collections leadership, curators, and managers, university administrators, public and private funders, and the scientific communities that use collections, among other entities whose perspectives and needs are important to the future vitality of biological collections. Currently, there are no shared mechanisms, meeting spaces, or virtual platforms that bring together all of these relevant and interested parties. Because biological collections are used in many disciplines for a multitude of research endeavors, the diversity of applications, objectives, funding agencies, and institutions involved amplifies the challenge of coordinating efforts, but it also provides opportunities for synthesis of information from multiple sources. Silos within the biological collections community exist, particularly in terms of the discipline represented, information-sharing, curatorial activities, and even funding opportunities, resulting in duplicated effort in some cases and, in other cases, parts of the collections community that have been seemingly left behind. Many current working

[4] See https://www.nceas.ucsb.edu.
[5] See https://www.sesync.org.
[6] See https://www.usanpn.org/usa-national-phenology-network.
[7] See https://www.astc.org.
[8] See https://www.usgs.gov/centers/powell-ctr/science.

groups and professional organizations[9] are engaged in parallel discussions, but sometimes these also lead to disconnected efforts, despite the many shared needs across all types of biological collections. An Action Center for Biological Collections could help streamline those efforts by fostering partnerships and promoting complementary activities.

Efforts to digitize the nation's biological collections have become a driving force for unity. In addition, advancements in cyberinfrastructure have increased our ability and extent to participate virtually to research and education events. For example, the iPlant Collaborative (Goff et al., 2011) or EarthCube,[10] both funded by NSF, create a virtual platform for their communities that combines research innovation with computing resources. Integrating virtual participation into a biological collection action center could promote productive spaces for interdisciplinary interactions, as biological specimens and associated data are increasingly accessed and used in a diverse array of research initiatives. As described in Chapter 5, the Integrated Digitized Biocollections and its Thematic Collections Networks, through funding from NSF, and to some extent biological resource centers, have provided some mechanisms for connecting the biological collections community through virtual training sessions, webinars, and a variety of other activities. Shared databases (e.g., Arctos, Symbiota) provide yet another vehicle for virtual cross-institutional interactions.

Research Coordination Networks (RCNs) funded by NSF (e.g., RCN award #1534564: A community of ex situ microbial germplasm collections in 2015; Biodiversity Collections Network) also serve to bring the collections community together, but generally only for the duration of the award. The activities of these, and other, previously funded RCNs provide a strong framework for the establishment of an Action Center for Biological Collections. Several professional societies have made large strides toward bringing biological collections personnel together, developing working groups to target a wide variety of needs. For example, the International Society for Biological and Environmental Repositories has worked to establish best practices and guidelines for maintaining the quality of biological repositories around the world, the American Phytopathological Society has been active in promoting culture collection support, and the Society for the Preservation of Natural History Collections has worked toward organizing a broad sector of biological collections personnel, primarily focusing on biological collection managers. These efforts are all positive steps, but strategic coordination across collections of all types is needed to ensure that the potential societal benefits of this vast resource are met. A biological collections–focused action center could facilitate training and further build and nurture communities of practice for research, education, workforce training, evaluation, and business strategies, among other needs. While institutions that curate, maintain, and use biological collections may have differing missions and sizes, they all face

[9] Such as the Interagency Working Group on Scientific Collections, One World Collection, World Federation for Culture Collections, Integrated Digitized Biocollections, Natural Science Collections Alliance, Society for the Preservation of Natural History Collections, Entomological Collections Network, and Society of Herbarium Curators, among others.

[10] See https://www.earthcube.org.

a complex balancing act to adapt to the evolving needs of science, education, and society. The coordinated action of a unified biological collections community could be a powerful resource that provides guidance, training, and support across a range of issues covered in this report such as, but not limited to:

- creating a national collections registry;
- engaging new user communities, including small collections;
- developing an evaluation plan and synthesizing quantitative and qualitative metrics;
- establishing a workforce pipeline for personnel;
- future-proofing financial models;
- sharing best practices and standards for quality control; and
- building a shared cyberinfrastructure.

Coordination and collaboration could bring biological collections of all sizes, all taxa, non-federal and federal, living stocks, and natural history together to establish shared leadership, vision, and strategic planning.

Coordination and sharing of knowledge will be critical for the biological collections community to be able to meet current and future needs and address the dynamic challenges of society and rapid global change (e.g., Cook et al., 2020). Biological collections play an important role in this endeavor, and the broader community has much to share and learn from one another. The nation's biological collections will be much more effective at meeting future societal needs if the community works together under coordinated leadership, vision, and strategy. The biological collections community needs an inclusive, integrated platform to strengthen the position of biological collections as a unified scientific infrastructure for the nation over the next decade and beyond. A national collections-focused action center dedicated to the support and use of biological collections could fill this need.

A National Decadal Survey for Biological Collections

Once a physical and virtual synthesis space to facilitate coordination and collaboration is created, this action center could facilitate the development and implementation of a national vision for research, education, and service to the nation in general. Many scientific communities work together to set priority research topics and the building of infrastructure needed to accomplish those priorities. Examples include the decadal surveys carried out by ocean science, astronomy, Earth science, planetary science, and materials science communities, which serve not only to unify the communities around a set of common goals but also to inform internal strategic planning of federal science funding agencies (NASEM, 2015, 2017b, 2018a, 2019b). A biological collections–focused decadal survey would establish a set of priorities that could only be accomplished with a concerted effort of the collective, rather than any one individual biological collection (e.g., an "Earthshot" effort aimed at revealing the three-dimensional morphology, associated genomes, and potential biotic interactions of all diversity on Earth). A decadal survey for the biological collections community will need to involve the natural history and living stocks collections communities. As evidenced by this report, both groups have particular needs and strengths that do not entirely overlap, so deeper coordination or understanding of the differences between the two in terms of strategy and planning will be mutually beneficial. The two communities often hold specimens derived from the same original gathering or isolation, and the digital linking of these separate parts will be greatly facilitated by a closer working relationship between the institutions holding material of common origin. The planning process would also need engagement across NSF directorates and programs to include a broader group of end users and stakeholders for biological collections. As recommended in the previous chapters, the collections community needs to make stronger connections with computer science, engineering, educational researchers, social science, and other disciplines not traditionally associated with biological collections, but that are becoming increasingly engaged users of biological collections. Cross-directorate participation in a decadal survey would help to strengthen these connections.

Such a visioning process would also benefit by reaching across federal agencies that support the biological collections infrastructure to develop plans for federal versus non-federal collections. Living stocks collections exemplify how complex the funding and end-user base of collections can be from NSF to the Department of Agriculture to the National Institutes of Health, and from traditional research conducted at universities to for-profit companies using living stocks collections to develop new medicines, vaccines, or crops. The artificial silos that inhibit collaborative action of funding agencies to support biological collections are not beneficial to science, research, or education moving forward in the United States. The most exciting and novel types of questions that can be answered using biological collections, the ones that potentially have the most benefit to society, can transcend disciplinary silos, funding agencies, and the nonprofit and for-profit world. Such partnerships can leverage resources and maximize progress and are expected to foster large, transdisciplinary programs that address complex, high-priority questions related to global change and public health. Such partnerships can maximize the value of both research and infrastructure investments and could help distribute the costs of biological collections infrastructure beyond the NSF Division of Biological Infrastructure.

Through broad discussion with the growing set of users and stakeholders, a decadal plan for biological collections could be developed. Such a plan could guide the development and expansion of the nation's biological collections, and become an important tool to share and leverage these resources. In addition, a potential eleventh Big Idea on understanding the sixth extinction (Ripple et al., 2020), will require robust national biological collections infrastructure as transdisciplinary collaborations focus on the breadth and implications of massive biodiversity loss. Working more broadly across the sciences and technology on such issues would help further integrate the biological collections community into research collaborations in more interesting and novel ways.

RECOMMENDATIONS FOR THE NEXT STEPS

Recommendation 8-1: The National Science Foundation, in collaboration with other institutions that provide funding and other types of support for biological collections, should help establish a permanent national Action Center for Biological Collections to coordinate action and knowledge, resources, and data-sharing among the nation's biological collections as they strive to meet the complex and often unpredictable needs of science and society. Such an action center should include a physical space and cyber-infrastructure to develop and implement collaborative strategic efforts and further build and nurture communities of practice for research, education, workforce training, evaluation, and business model development, among other community-wide needs.

Recommendation 8-2: The National Science Foundation should lead efforts to develop a vision and strategy, such as a decadal survey, for targeted growth of the nation's biological collections, their infrastructure, and their ability to serve a broader range of users and scientific and educational needs. The vision and strategy should take into consideration the diverse capabilities and needs of all types of collections and diverse array of end users, and set long-range priorities that could only be accomplished with a concerted, collaborative effort of the nation's biological collections.

Recommendation 8-3: The National Science Foundation (NSF) Directorate for Biological Sciences should expand its partnership capabilities more broadly across NSF, other federal agencies, international programs, and other sectors. Such partnerships can maximize investments in support of a national Action Center for Biological Collections, and the development of a national vision and strategy and help spread the cost of such major endeavors beyond the NSF Directorate for Biological Sciences.

References

AAAS (American Association for the Advancement of Science). 2015. *Vision and change in undergraduate biology education: Chronicling change, inspiring the future*. Washington, DC: American Association for the Advancement of Science.

Abbas, C. A. 2006. Production of antioxidants, aromas, colours, flavours, and vitamins by yeasts. In A. Querol and G. Fleet (eds.), *Yeasts in food and beverages*. Berlin, Germany: Springer.

Adam, P. S., G. Borrel, C. Brochier-Armanet, and S. Gribaldo. 2017. The growing tree of archaea: New perspectives on their diversity, evolution and ecology. *ISME Journal* 11(11):2407–2425.

Aguirre-Liguori, J. A., S. Ramirez Barahona, P. Tiffin, and L. E. Eguiarte. 2019. Climate change is predicted to disrupt patterns of local adaptation in wild and cultivated maize. *Proceedings of the Royal Society B: Biological Sciences* 286:20190486.

AIC (American Institute for Conservation). 2013. *Collection Care Network*. https://www.culturalheritage.org/docs/default-source/publications/reports/collection-care-staff-survey-report.pdf?sfvrsn=8 (accessed August 24, 2020).

Allmon, W. D. 1994. The value of natural history collections. *Curator: The Museum Journal* 37(2):83–89.

Alverdy, J. C., and E. B. Chang. 2008. The re-emerging role of the intestinal microflora in critical illness and inflammation: Why the gut hypothesis of sepsis syndrome will not go away. *Journal of Leukocyte Biology* 83(3):461–466.

Anderson, J. F., C. R. Vossbrinck, T. G. Andreadis, A. Iton, W. H. Beckwith, 3rd, and D. R. Mayo. 2001. A phylogenetic approach to following West Nile virus in Connecticut. *Proceedings of the National Academy of Sciences* 98(23):12885–12889.

Anderson, R. P. 2012. Harnessing the world's biodiversity data: Promise and peril in ecological niche modeling of species distributions. *Annals of the New York Academy of Sciences* 1260:66–80.

Ankeny, R. A. 2019. *A philosophical perspective on biological collections*. Presentation to Committee on Biological Collections: Their Past, Present, and Future Contributions and Options for Sustaining Them by Webinar on February 15, 2019. https://vimeo.com/326662039 (accessed August 21, 2020).

Antes, L. A., A. Mart, and J. M. DuBois. 2016. Are leadership and management essential for good research? An interview study of genetic researchers. *Journal of Empirical Research on Human Research Ethics* 11(5):408–423.

Antonelli, A., M. Ariza, J. Albert, T. Andermann, J. Azevedo, C. Bacon, S. Faurby, T. Guedes, C. Hoorn, L. G. Lohmann, P. Matos-Maravi, C. D. Ritter, I. Sanmartin, D. Silvestro, M. Tejedor, H. Ter Steege, H. Tuomisto, F. P. Werneck, A. Zizka, and S. V. Edwards. 2018. Conceptual and empirical advances in neotropical biodiversity research. *PeerJ* 6:e5644.

Antunes, A., E. Stackebrandt, and N. Lima. 2016. Fueling the bio-economy: European culture collections and microbiology education and training. *Trends in Microbiology* 24(2):77–79.

Araujo, A. L. 2019. The death of Brazil's national museum. *American Historical Review* 124(2):569–580.

Arbeláez-Cortés, E., A. R. Acosta-Galvis, C. DoNascimiento, D. Espitia-Reina, A. González-Alvarado, and C. A. Medina. 2017. Knowledge linked to museum specimen vouchers: Measuring scientific production from a major biological collection in Colombia. *Scientometrics* 112(3):1323–1341.

Archer, L., E. Dawson, J. DeWitt, A. Seakins, and B. Wong. 2015. "Science capital": A conceptual, methodological, and empirical argument for extending Bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching* 52(7):922–948.

Austin, A. E. 2018. *Vision and change in undergraduate biology education: Unpacking a movement and sharing lessons learned*. Washington, DC: American Association for the Advancement of Science.

Autumn, K., M. Sitti, Y. A. Liang, A. M. Peattie, W. R. Hansen, S. Sponberg, T. W. Kenny, R. Fearing, J. N. Israelachvili, and R. J. Full. 2002. Evidence for van der Waals adhesion in gecko setae. *Proceedings of the National Academy of Sciences* 99(19):12252–12256.

Autumn, K., P. H. Niewiarowski, and J. B. Puthoff. 2014. Gecko adhesion as a model system for integrative biology, interdisciplinary science, and bioinspired engineering. *Annual Review of Ecology, Evolution, and Systematics* 45(1):445–470.

Baker, R. J., L. C. Bradley, R. D. Bradley, and H. J. Garner. 2014. *"Door to drawer" costs of curation, installation, documentation, databasing, and long-term care of mammal voucher specimens in natural history collections*. Occasional Papers no. 323. Lubbock, TX: Museum of Texas Tech University.

Bakker, F. T., A. Antonelli, J. A. Clarke, J. A. Cook, S. V. Edwards, P. G. P. Ericson, S. Faurby, N. Ferrand, M. Gelang, R. G. Gillespie, M. Irestedt, K. Lundin, E. Larsson, P. Matos-Maraví, J. Müller, T. von Proschwitz, G. K. Roderick, A. Schliep, N. Wahlberg, J. Wiedenhoeft, and M. Källersjö. 2020. The global museum: Natural history collections and the future of evolutionary science and public education. *PeerJ* 8:e8225.

Balengée, B., and N. Triscott. 2010. *Malamp: The occurrence of deformities in amphibians*. Sheffield City Centre, Sheffield: Arts Catalyst and Yorkshire Sculpture Park.

Ball-Damerow, J. E., L. Brenskelle, N. Barve, P. S. Soltis, P. Sierwald, R. Bieler, R. LaFrance, A. H. Ariño, and R. Guralnick. 2019. Research applications of primary biodiversity databases in the digital age. *PLOS ONE* 14(9).

Ballard, H. L., L. D. Robinson, A. N. Young, G. B. Pauly, L. M. Higgins, R. F. Johnson, and J. C. Tweddle. 2017. Contributions to conservation outcomes by natural history museum-led citizen science: Examining evidence and next steps. *Biological Conservation* 208:87–97.

Ballard, H. L., E. M. Harris, and C. G. H. Dixon. 2018. *Science identity and agency in community and citizen science: Evidence and potential.* Paper commissioned for the Committee on Designing Citizen Science to Support Science Learning, Board on Science Education, National Academy of Sciences, Engineering, and Medicine. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_189606.pdf (accessed September 6, 2020).

Baron, J. S., A. Specht, E. Garnier, P. Bishop, C. A. Campbell, F. W. Davis, B. Fady, D. Field, L. J. Gross, S. M. Guru, B. S. Halpern, S. E. Hampton, P. R. Leavitt, T. R. Meagher, J. Ometto, J. N. Parker, R. Price, C. H. Rawson, A. Rodrigo, L. A. Sheble, and M. Winter. 2017. Synthesis centers as critical research infrastructure. *BioScience* 67(8):750–759.

Bates, J. 2007. Natural history museums: World centers of biodiversity knowledge now and in the future. *The Systematist* 29:3–6.

Bauerle, C., A. DePass, D. Lynn, C. O'Connor, S. Singer, M. Withers, C. W. Anderson, S. Donovan, S. Drew, D. Ebert-May, L. Gross, S. G. Hoskins, J. Labov, D. Lopatto, W. McClatchey, P. Varma-Nelson, N. Pelaez, M. Poston, K. Tanner, D. Wessner, H. White, W. Wood, and D. Wubah. 2011. *Vision and change in undergraduate biology education: A call to action*. Washington, DC: American Association for the Advancement of Science.

BCoN (Biodiversity Collections Network). 2018. *Integration, attribution, and value in the web of Natural History Museum data: A needs assessment workshop*. February 13–14, Lawrence, KS. https://bcon.aibs.org/wp-content/uploads/2018/05/BCoN-Needs-Assessment-workshop-report-1.pdf (accessed November 30, 2020).

Bebber, D. P., M. A. Carine, J. R. I. Wood, A. H. Wortley, D. J. Harris, G. T. Prance, G. Davidse, J. Paige, T. D. Pennington, N. K. B. Robson, and R. W. Scotland. 2010. Herbaria are a major frontier for species discovery. *Proceedings of the National Academy of Sciences* 107(51):22169–22171.

Becker, P. R., and S. A. Wise. 2006. The U.S. National Biomonitoring Specimen Bank and the Marine Environmental Specimen Bank. *Journal of Environmental Monitoring* 8(8):795–799.

Becker, P., M. Bosschaerts, P. Chaerle, H.-M. Daniel, A. Hellemans, A. Olbrechts, L. Rigouts, A. Wilmotte, and M. Hendrickx. 2019. Public microbial resources centres: Key hubs for fair microorganisms and genetic materials. *Applied and Environmental Microbiology* 85(21):e01444-19.

Beckmann, E., G. Estavillo, U. Mathesius, M. Djordjevic, and A. Nicotra. 2015. The plant detectives: Innovative undergraduate teaching to inspire the next generation of plant biologists. *Frontiers in Plant Science* 6:729.

Belitz, M. W., L. K. Hendrick, M. J. Monfils, D. L. Cuthrell, C. J. Marshall, A. Y. Kawahara, N. S. Cobb, J. M. Zaspel, A. M. Horton, S. L. Huber, A. D. Warren, G. A. Forthaus, and A. K. Monfils. 2018. Aggregated occurrence records of the federally endangered poweshiek skipperling (*Oarisma poweshiek*). *Biodiversity Data Journal* 6:e29081.

Bellen, H. J., C. Tong, and H. Tsuda. 2010. 100 years of *Drosophila* research and its impact on vertebrate neuroscience: A history lesson for the future. *Nature Reviews Neuroscience* 11(7):514–522.

Bernard, R. E., and E. H. G. Cooperdock. 2018. No progress on diversity in 40 years. *Nature Geoscience* 11:292–295.

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. http://www.obfs.org/fsml-future (accessed August 24, 2020).

Bloom, T. D. S., A. Flower, and E. G. DeChaine. 2017. Why georeferencing matters: Introducing a practical protocol to prepare species occurrence records for spatial analysis. *Ecology and Evolution* 8(1):765–777.

Bouadjenek, M. R., J. Zobel, and K. Verspoor. 2019. Automated assessment of biological database assertions using the scientific literature. *BMC Bioinformatics* 20(1):216.

Boundy-Mills, K. L., E. Glantschnig, I. N. Roberts, A. Yurkov, S. Casaregola, H.-M. Daniel, M. Groenewald, and B. Turchetti. 2016. Yeast culture collections in the twenty-first century: New opportunities and challenges. *Yeast* 33(7):243–260.

Boundy-Mills, K., K. McCluskey, P. Elia, J. A. Glaeser, D. L. Lindner, D. R. Nobles, Jr., J. Normanly, F. M. Ochoa-Corona, J. A. Scott, T. J. Ward, K. M. Webb, K. Webster, and J. E. Wertz. 2019. Preserving U.S. microbe collections sparks future discoveries. *Journal of Applied Microbiology* 129(2):162–174.

Bozeman, B., D. Fay, and C. P. Slade. 2013. Research collaboration in universities and academic entrepreneurship: The state of the art. *Journal of Technology Transfer* 38(1):1–67.

Bradley, R. D. 2012. *Cost of collecting and preparing mammal voucher specimens for natural history collections*. Occasional Papers no. 313. Lubbock, TX: Museum of Texas Tech University.

Brenskelle, L., R. P. Guralnick, M. Denslow, and B. J. Stucky. 2020. Maximizing human effort for analyzing scientific images: A case study using digitized herbarium sheets. *Applications in Plant Sciences* 8(6):e11370.

Brock, T. D. 1967. Life at high temperatures. *Science* 158(3804):1012–1019.

Brock, T. D., and M. L. Brock. 1967. The measurement of chlorophyll, primary productivity, photophosphorylation, and macromolecules in benthic algal mats. *Limnology and Oceanography* 12(4):600–605.

Brock, T. D., and M. L. Brock. 1968. Measurement of steady-state growth rates of a thermophilic alga directly in nature. *Journal of Bacteriology* 95(3):811–815.

Brock, T. D., and H. Freeze. 1969. *Thermus aquaticus* gen. N. and sp. N., a nonsporulating extreme thermophile. *Journal of Bacteriology* 98(1):289–297.

Brooks, D. R., E. P. Hoberg, and W. A. Boeger. 2019. *The Stockholm paradigm: Climate change and emerging disease*. Chicago, IL: University of Chicago Press.

Burgin, C. J., J. P. Colella, P. L. Kahn, and N. S. Upham. 2018. How many species of mammals are there? *Journal of Mammalogy* 99(1):1–14.

Cai, L., B. B. Jorgensen, C. A. Suttle, M. He, B. A. Cragg, N. Jiao, and R. Zhang. 2019. Active and diverse viruses persist in the deep sub-seafloor sediments over thousands of years. *ISME Journal* 13(7):1857–1864.

Cameron, D. E., C. J. Bashor, and J. J. Collins. 2014. A brief history of synthetic biology. *Nature Reviews Microbiology* 12(5):381–390.

Campbell, L. G., S. Mehtani, M. E. Dozier, and J. Rinehart. 2013. Gender heterogeneous working groups produce higher quality science. *PLOS ONE* 8:e79147.

Campbell, L. D., J. J. Astrin, R. Brody, Y. D. Souza, J. G. Giri, A. A. Patel, M. Rawley-Payne, A. Rush, and N. Sieffert. 2018. The 2018 revision of the ISBER best practices: Summary of changes and the editorial team's development process. *Biopreservation and Biobanking* 16:3–6.

Campos, M., S. K. Govers, I. Irnov, G. S. Dobihal, F. Cornet, and C. Jacobs-Wagner. 2018. Genome-wide phenotypic analysis of growth, cell morphogenesis, and cell cycle events in *Escherichia coli*. *Molecular Systems Biology* 14(6):e7573.

Carranza-Rojas, J., H. Goeau, P. Bonnet, E. Mata-Montero, and A. Joly. 2017. Going deeper in the automated identification of herbarium specimens. *BMC Evolutionary Biology* 17(1):181.

Catlin-Groves, C. L. 2012. The citizen science landscape: From volunteers to citizen sensors and beyond. *International Journal of Zoology* 2012:349630.

Cavallari, D. C., R. B. Salvador, and B. Rodrigues Da Cunha. 2014. Dangers to malacological collections: Bynesian decay and pyrite decay. *Collection Forum* 28(1–2):35–46.

Ceballos, G., P. R. Ehrlich, and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences* 114(30):E6089.

Center for Health Workforce Studies and NASW (National Association of Social Workers) Center for Workforce Studies. 2006. *Licensed social workers in the United States, 2004*. Washington, DC: National Association of Social Workers. https://www.socialworkers.org/LinkClick.aspx?fileticket=r4K7DWvfskk%3D&portalid=0 (accessed September 6, 2020).

Chandras, C., T. Weaver, M. Zouberakis, D. Smedley, K. Schughart, N. Rosenthal, J. M. Hancock, G. Kollias, P. N. Schofield, and V. Aidinis. 2009. Models for financial sustainability of biological databases and resources. *Database* 2009:bap017.

Chapman, A., and J. Wieczorek. 2006. *Guide to best practices for georeferencing*. Copenhagen, Denmark: Global Biodiversity Information Facility.

Chapman, A. D., L. Belbin, P. F. Zermoglio, J. Wieczorek, P. J. Morris, M. Nicholls, E. R. Rees, A. K. Veiga, A. Thompson, A. M. Saraiva, S. A. James, C. Gendreau, A. Benson, and D. Schigel. 2020. Developing standards for improved data quality and for selecting fit for use biodiversity data. *Biodiversity Information Science and Standards* 4:e50889.

Chen, M. L., A. Doddi, J. Royer, L. Freschi, M. Schito, M. Ezewudo, I. S. Kohane, A. Beam, and M. Farhat. 2019. Beyond multidrug resistance: Leveraging rare variants with machine and statistical learning models in *Mycobacterium tuberculosis* resistance prediction. *EBioMedicine* 43:356–369.

Clark, J. A., J. M. Hoekstra, P. D. Boersma, and P. Kareiva. 2002. Improving U.S. Endangered Species Act recovery plans: Key findings and recommendations of the SCB Recovery Plan Project. *Conservation Biology* 16(6):1510–1519.

Cobb, N. S., L. F. Gall, J. M. Zaspel, N. J. Dowdy, L. M. McCabe, and A. Y. Kawahara. 2019. Assessment of North American arthropod collections: Prospects and challenges for addressing biodiversity research. *PeerJ* 7:e8086.

Colella, J. P., S. L. Talbot, C. Brochmann, E. B. Taylor, E. P. Hoberg, and J. A. Cook. 2020. Conservation genomics in a changing Arctic. *Trends in Ecology & Evolution* 35(2):149–162.

Cook, J. A., C. Brochmann, S. L. Talbot, V. Fedorov, E. B. Taylor, R. Väinölä, E. P. Hoberg, M. Kholodova, and K. P. Magnusson. 2013. Genetic perspectives on Arctic biodiversity. In H. Meltofte (ed.), *Arctic biodiversity assessment: Status and trends in Arctic biodiversity*. Akureyri, Iceland: Conservation of Arctic Fauna and Flora Committee. Pp. 459–483.

Cook, J. A., S. V. Edwards, E. A. Lacey, R. P. Guralnick, P. S. Soltis, D. E. Soltis, C. K. Welch, K. C. Bell, K. E. Galbreath, C. Himes, J. M. Allen, T. A. Heath, A. C. Carnaval, K. L. Cooper, M. Liu, J. Hanken, and S. Ickert-Bond. 2014. Natural history collections as emerging resources for innovative education. *BioScience* 64(8):725–734.

Cook, J. A., K. E. Galbreath, K. C. Bell, M. L. Campbell, S. Carrière, J. P. Colella, N. G. Dawson, J. L. Dunnum, R. P. Eckerlin, V. Fedorov, S. E. Greiman, G. M. S. Haas, V. Haukisalmi, H. Henttonen, A. G. Hope, D. Jackson, T. S. Jung, A. V. Koehler, J. M. Kinsella, D. Krejsa, S. J. Kutz, S. Liphardt, S. O. MacDonald, J. L. Malaney, A. Makarikov, J. Martin, B. S. McLean, R. Mulders, B. Nyamsuren, S. L. Talbot, V. V. Tkach, A. Tsvetkova, H. M. Toman, E. C. Waltari, J. S. Whitman, and E. P. Hoberg. 2016. 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(3):585–617.

Cook, J. A., S. Arai, B. Armién, J. Bates, C. A. C. Bonilla, M. B. d. S. Cortez, J. L. Dunnum, A. W. Ferguson, K. M. Johnson, F. A. A. Khan, D. L. Paul, D. M. Reeder, M. A. Revelez, N. B. Simmons, B. M. Thiers, C. W. Thompson, N. S. Upham, M. P. M. Vanhove, P. W. Webala, M. Weksler, R. Yanagihara, and P. S. Soltis. 2020. Integrating biodiversity infrastructure into pathogen discovery and mitigation of emerging infectious diseases. *BioScience* 70(6):531–534.

Coutinho, F. H., R. Rosselli, and F. Rodriguez-Valera. 2019. Trends of microdiversity reveal depth-dependent evolutionary strategies of viruses in the Mediterranean. *mSystems* 4(6):e00554-19.

Crane, K. 2019. *The costs and value of federal scientific collections*. Presentation to Committee on Biological Collections: Their Past, Present, and Future Contributions and Options for Sustaining Them by Webinar on July 9, 2019. https://vimeo.com/348902417 (accessed August 21, 2020).

Cross, K. L., J. H. Campbell, M. Balachandran, A. G. Campbell, S. J. Cooper, A. Griffen, M. Heaton, S. Joshi, D. Klingeman, E. Leys, Z. Yang, J. M. Parks, and M. Podar. 2019. Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. *Nature Biotechnology* 37:1314–1321.

Cunningham, A. A., P. Daszak, and J. L. N. Wood. 2017. One Health, emerging infectious diseases and wildlife: Two decades of progress? *Philosophical Transactions of the Royal Society B: Biological Sciences* 372(1725):20160167.

Dalton, R. 2003. Natural history collections in crisis as funding is slashed. *Nature* 423:575.

Dance, A. 2017. How museum work can combine research and public engagement. *Nature* 552:279–281.

Daru, B. H., D. S. Park, R. B. Primack, C. G. Willis, D. S. Barrington, T. J. S. Whitfeld, T. G. Seidler, P. W. Sweeney, D. R. Foster, A. M. Ellison, and C. C. Davis. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist* 217(2):939–955.

Dawes, L. 2016. *Talk box: Speaking and listening activities for learning at key stage 1*. New York: Routledge.

De Vero, L., M. B. Boniotti, M. Budroni, P. Buzzini, S. Cassanelli, R. Comunian, M. Gullo, A. F. Logrieco, I. Mannazzu, R. Musumeci, I. Perugini, G. Perrone, A. Pulvirenti, P. Romano, B. Turchetti, and G. C. Varese. 2019. Preservation, characterization and exploitation of microbial biodiversity: The perspective of the Italian Network of Culture Collections. *Microorganisms* 7(12):685.

Deng, B. 2015. Plant collections left in the cold by cuts. *Nature* 523(7558):16.

Di Marco, M., S. Chapman, G. Althor, S. Kearney, C. Besancon, N. Butt, J. M. Maina, H. P. Possingham, K. Rogalla von Bieberstein, O. Venter, and J. E. M. Watson. 2017. Changing trends and persisting biases in three decades of conservation science. *Global Ecology and Conservation* 10:32–42.

Diah, S. Z. M., S. B. Karman, and I. C. Gebeshuber. 2014. Nanostructural colouration in Malaysian plants: Lessons for biomimetics and biomaterials. *Journal of Nanomaterials* 2014(2):1–15.

Díaz, S., J. Settele, E. Brondízio, H. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. Brauman, S. Butchart, K. Chan, L. Garibaldi, K. Ichii, J. Liu, S. M. Subrmanian, G. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. R. Chowdhury, Y.-J. Shin, I. Visseren-Hamakers, K. Wilis, and C. Zayas. 2019. *Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Bonn, Germany: Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.

DiEuliis, D., K. R. Johnson, S. S. Morse, and D. E. Schindel. 2016. Opinion: Specimen collections should have a much bigger role in infectious disease research and response. *Proceedings of the National Academy of Sciences* 113(1):4–7.

Disis, M. L., and J. T. Slattery. 2010. The road we must take: Multidisciplinary team science. *Science Translational Medicine* 2(2):22cm9.

Drew, J. A., C. S. Moreau, and M. L. J. Stiassny. 2017. Digitization of museum collections holds the potential to enhance researcher diversity. *Nature Ecology and Evolution* 1(12):1789–1790.

DuBay, S. G., and C. C. Fuldner. 2017. Bird specimens track 135 years of atmospheric black carbon and environmental policy. *Proceedings of the National Academy of Sciences* 114(43):11321–11326.

Dunnum, J. L., R. Yanagihara, K. M. Johnson, B. Armien, N. Batsaikhan, L. Morgan, and J. A. Cook. 2017. Biospecimen repositories and integrated databases as critical infrastructure for pathogen discovery and pathobiology research. *PLOS Neglected Tropical Diseases* 11(1):e0005133.

Dunnum, J. L., B. S. McLean, and R. C. Dowler. 2018. Mammal collections of the Western Hemisphere: A survey and directory of collections. *Journal of Mammalogy* 99:1307–1322.

Durden, L. A., J. E. Keirans, and J. H. Oliver, Jr. 1996. The U.S. National Tick Collection: A vital resource for systematics and human and animal welfare. *American Entomologist* 42(4):239–243.

Dutt, K. 2020. Race and racism in the geosciences. *Nature Geoscience* 13:2–3.

Edwards, S. V., S. Birks, R. T. Brumfield, and R. Hanner. 2005. Future of avian genetic resources collections: Archives of evolutionary and environmental history. *Auk* 122:979–984.

Elkin, L., and C. A. Norris (eds.). 2019. *Preventive conservation: Collection storage*. New York: Society for the Preservation of Natural History Collections.

Elkins, J. T., and N. M. L. Elkins. 2007. Teaching geology in the field: Significant geoscience concept gains in entirely field-based introductory geology courses. *Journal of Geoscience Education* 55(2):126–132.

Elliott, M. J., J. H. Poelen, and J. A. B. Fortes. 2020. Toward reliable biodiversity dataset references. *Ecological Informatics* 59:101132.

Ellison, C. E., C. Hall, D. Kowbel, J. Welch, R. B. Brem, N. L. Glass, and J. W. Taylor. 2011. Population genomics and local adaptation in wild isolates of a model microbial eukaryote. *Proceedings of the National Academy of Sciences* 108(7):2831–2836.

Ellwood, E. R., B. A. Dunkel, P. Flemons, R. Guralnick, G. Nelson, G. Newman, S. Newman, D. Paul, G. Riccardi, N. Rios, K. C. Seltmann, and A. R. Mast. 2015. Accelerating the digitization of biodiversity research specimens through online public participation. *BioScience* 65(4):383–396.

Ellwood, E. R., P. Kimberly, R. Guralnick, P. Flemons, K. Love, S. Ellis, J. M. Allen, J. H. Best, R. Carter, S. Chagnoux, R. Costello, M. W. Denslow, B. A. Dunckel, M. M. Ferriter, E. E. Gilbert, C. Goforth, Q. Groom, E. R. Krimmel, R. LaFrance, J. L. Martinec, A. N. Miller, J. Minnaert-Grote, T. Nash, P. Oboyski, D. L. Paul, K. D. Pearson, N. D. Pentcheff, M. A. Roberts, C. E. Seltzer, P. S. Soltis, R. Stephens, P. W. Sweeney, M. von Konrat, A. Wall, R. Wetzer, C. Zimmerman, and A. R. Mast. 2018. Worldwide engagement for digitizing biocollections (WeDigBio): The biocollections community's citizen-science space on the calendar. *BioScience* 68(2):112–124.

Ellwood, E., A. Monfils, L. White, D. Linton, N. Douglas, and M. Phillips. 2019. Developing a data-literate workforce through BLUE: Biodiversity Literacy in Undergraduate Education. *Biodiversity Information Science and Standards* 3:e37339.

EPA (Environmental Protection Agency). 2002. *National Water Quality Inventory 2000 report*. Washington, DC: EPA Office of Water.

Falk, J. H., and L. D. Dierking. 2013. *The museum experience revisited, 2nd ed*. New York: Routledge.

Falk, J. H., and L. D. Dierking. 2018. *Learning from museums, 2nd ed*. Lanham, MD: Rowman & Littlefield.

Fanning, T. G., R. D. Slemons, A. H. Reid, T. A. Janczewski, J. Dean, and J. K. Taubenberger. 2002. 1917 avian influenza virus sequences suggest that the 1918 pandemic virus did not acquire its hemagglutinin directly from birds. *Journal of Virology* 76(15):7860–7862.

Ferguson, N. M., A. P. Galvani, and R. M. Bush. 2003. Ecological and immunological determinants of influenza evolution. *Nature* 422(6930):428–433.

Flaherty, C. 2017. Louisiana-Monroe natural history collections are safe. *Inside Higher Ed*, July 17. https://www.insidehighered.com/quicktakes/2017/07/05/louisiana-monroe-natural-history-collections-are-safe (accessed November 30, 2020).

Flattau, P. E., M. Boeckmann, R. de la Cruz, P. Lagasse, N. Mitchell, M. Patterson, and D. Singpurwalla. 2007. *Scientific collections: Mission-critical infrastructure for federal scientific agencies*. Washington, DC: Science and Technology Policy Institute.

Fontaine, B., A. Perrard, and P. Bouchet. 2012. 21 years of shelf life between discovery and description of new species. *Current Biology* 22(22):R943–R944.

Ford-Lloyd, B. V., M. Schmidt, S. J. Armstrong, O. Barazani, J. Engels, R. Hadas, K. Hammer, S. P. Kell, D. Kang, K. Khoshbakht, Y. Li, C. Long, B.-R. Lu, K. Ma, V. T. Nguyen, L. Qiu, S. Ge, W. Wei, Z. Zhang, and N. Maxted. 2011. Crop wild relatives—undervalued, underutilized and under threat? *BioScience* 61(7):559–565.

Forster, J. D., H. Noh, S. F. Liew, V. Saranathan, C. F. Schreck, L. Yang, J. G. Park, R. O. Prum, S. G. J. Mochrie, C. S. O'Hern, H. Cao, and E. R. Dufresne. 2010. Biomimetic isotropic nanostructures for structural coloration. *Advanced Materials* 22(26–27):2939–2944.

Freedman, L. P., I. M. Cockburn, and T. S. Simcoe. 2015. The economics of reproducibility in preclinical research. *PLOS Biology* 13(6):e1002165.

Freeman, R. B., and W. Huang. 2015. Collaborating with people like me: Ethnic coauthorship within the United States. *Journal of Labor Economics* 33:S289–S318.

Freeze, H., and T. D. Brock. 1970. Thermostable aldolase from *Thermus aquaticus*. *Journal of Bacteriology* 101(2):541–550.

Friedman. A. 2007. *A framework for evaluating impacts of informal science education projects*. Washington, DC: National Science Foundation. https://www.informalscience.org/framework-evaluating-impacts-informal-science-education-projects (accessed November 18, 2020).

Furman, J. L., and S. Stern. 2011. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review* 101:1933–1963.

Galbreath, K. E., E. P. Hoberg, J. A. Cook, B. Armien, K. C. Bell, M. L. Campbell, J. L. Dunnum, A. T. Dursahinhan, R. P. Eckerlin, S. L. Gardner, S. E. Greiman, H. Henttonen, F. A. Jimenez, A. V. A. Koehler, B. Nyamsuren, V. V. Tkach, F. Torres-Perez, A. Tsvetkova, and A. G. Hope. 2019. Building an integrated infrastructure for exploring biodiversity: Field collections and archives of mammals and parasites. *Journal of Mammalogy* 100(2):382–393.

Gelabert, P., M. Sandoval-Velasco, I. Olalde, R. Fregel, A. Rieux, R. Escosa, C. Aranda, K. Paaijmans, I. Mueller, M. T. Gilbert, and C. Lalueza-Fox. 2016. Mitochondrial DNA from the eradicated European *Plasmodium vivax* and *P. falciparum* from 70-year-old slides from the Ebro Delta in Spain. *Proceedings of the National Academy of Science* 113(41):11495–11500.

George, S. B. 2015. The "stuff" of museums: Collections, interactivity, and a call to action. *Dimensions*, July 20. https://www.astc.org/astc-dimensions/the-stuff-of-museums-collections-interactivity-and-a-call-to-action (accessed September 6, 2020).

George, S. B. 2019. *Key components of sustainable mission and infrastructure for a biological collection*. Presentation to Committee on Biological Collections: Their Past, Present, and Future Contributions and Options for Sustaining Them in Washington, DC, on February 7, 2019. https://vimeo.com/326425745 (accessed November 30, 2020).

Glass, G. E., T. M. Shields, R. R. Parmenter, D. Goade, J. N. Mills, J. Cheek, J. Cook, and T. L. Yates. 2006. Hantavirus risk in 2006 for U.S. Southwest. Occasional Papers no. 255. Lubbock, TX: Museum of Texas Tech University.

Goff, S., M. Vaughn, S. McKay, E. Lyons, A. Stapleton, D. Gessler, N. Matasci, L. Wang, M. Hanlon, A. Lenards, A. Muir, N. Merchant, S. Lowry, S. Mock, M. Helmke, A. Kubach, M. Narro, N. Hopkins, D. Micklos, U. Hilgert, M. Gonzales, C. Jordan, E. Skidmore, R. Dooley, J. Cazes, R. McLay, Z. Lu, S. Pasternak, L. Koesterke, W. Piel, R. Grene, C. Noutsos, K. Gendler, X. Feng, C. Tang, M. Lent, S.-j. Kim, K. Kvilekval, B. S. Manjunath, V. Tannen, A. Stamatakis, M. Sanderson, S. Welch, K. Cranston, P. Soltis, D. Soltis, B. O'Meara, C. Ane, T. Brutnell, D. Kleibenstein, J. White, J. Leebens-Mack, M. Donoghue, E. Spalding, T. Vision, C. Myers, D. Lowenthal, B. Enquist, B. Boyle, A. Akoglu, G. Andrews, S. Ram, D. Ware, L. Stein, and D. Stanzione. 2011. The iPlant collaborative: Cyberinfrastructure for plant biology. *Frontiers in Plant Science* 2:34.

Golembiewski, K. 2016. 5 behind-the-scenes specimens with links to Darwin. The Field Museum Blog, February 11. https://www.fieldmuseum.org/blog/5-behind-scenes-specimens-links-darwin (accessed November 30, 2020).

Goodwin, Z. A., D. J. Harris, D. Filer, J. R. Wood, and R. W. Scotland. 2015. Widespread mistaken identity in tropical plant collections. *Current Biology* 25(22):R1066–R1067.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution* 19(9):497–503.

Green, D. W., J. A. Watson, H. S. Jung, and G. S. Watson. 2019. Natural history collections as inspiration for technology. *Bioessays* 41(2):e1700238.

Grinnell Resurvey Project. n.d. *Home page*. University of California, The Museum of Vertebrate Zoology. http://mvz.berkeley.edu/Grinnell (accessed July 26, 2020).

Groom, Q. J., C. O'Reilly, and T. Humphrey. 2014. Herbarium specimens reveal the exchange network of British and Irish botanists, 1856–1932. *New Journal of Botany* 4(2):95–103.

Groom, Q., P. Desmet, L. Reyserhove, T. Adriaens, D. Oldoni, S. Vanderhoeven, S. J. Baskauf, A. Chapman, M. McGeoch, R. Walls, J. Wieczorek, J. R. U. Wilson, P. F. F. Zermoglio, and A. Simpson. 2019. Improving Darwin Core for research and management of alien species. *Biodiversity Information Science and Standards* 3:e38084.

Gropp, R. E. 2003. Are university natural science collections going extinct? *BioScience* 53:550.

Gropp, R. E. 2004. Budget cuts affecting natural history. *Science* 306:811.

Guarneros-Meza, V., J. Downe, and S. Martin. 2018. Defining, achieving, and evaluating collaborative outcomes: A theory of change approach. *Public Management Review* 20(10):1562–1580.

Güntsch, A., R. Hyam, G. Hagedorn, S. Chagnoux, D. Röpert, A. Casino, G. Droege, F. Glöckler, K. Gödderz, Q. Groom, J. Hoffmann, A. Holleman, M. Kempa, H. Koivula, K. Marhold, N. Nicolson, V. S. Smith, and D. Triebel. 2017. Actionable, long-term stable and semantic web compatible identifiers for access to biological collection objects. *Database* 2017:bax003.

Guralnick, R., T. Conlin, J. Deck, B. J. Stucky, and N. Cellinese. 2014. The trouble with triplets in biodiversity informatics: A data-driven case against current identifier practices. *PLOS ONE* 9(12):e114069.

Guralnick, R. P., P. F. Zermoglio, J. Wieczorek, R. LaFrance, D. Bloom, and L. Russell. 2016. The importance of digitized biocollections as a source of trait data and a new VertNet resource. *Database* 2016:baw158.

Gürel, E., and M. Tat. 2017. SWOT analysis: A theoretical review. *Journal of International Social Research* 10(51):994–1006.

Guthrie, S., W. Wamae, S. Diepeveen, S. Wooding, and J. Grant. 2013. *Measuring research: A guide to research evaluation frameworks and tools*. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/monographs/MG1217.html (accessed August 24, 2020).

Habig, B., P. Gupta, B. Levine, and J. Adams. 2018. An informal science education program's impact on STEM major and STEM career outcomes. *Research in Science Education* 50(3):1051–1074.

Hall, K. 1998. Storage concerns for geological collections. *Conserve O Gram* 11(2):1–4. https://www.nps.gov/museum/publications/conserveogram/11-02.pdf (accessed November 30, 2020).

Hao, M. J., and R. Yazdanifard. 2015. How effective leadership can facilitate change in organizations through improvement and innovation. *Global Journal of Management and Business Research: Administration and Managemen*t 15(9):1–5.

Harmon, A., D. T. J. Littlewood, and C. L. Wood. 2019. Parasites lost: Using natural history collections to track disease change across deep time. *Frontiers in Ecology and the Environment* 17(3):157–166.

Haston, E. M., R. W. N. Cubey, M. Pullan, H. Atkins, and D. Harris. 2012. Developing integrated workflows for the digitisation of herbarium specimens using a modular and scalable approach. *Zookeys* 209:93–102.

Hawks, C., M. McCann, K. Makos, L. Goldberg, D. Hinkamp, J. D. Ertel, and P. Silence (eds.). 2010. *Health and safety for museum professionals*. Washington, DC: Society for the Preservation of Natural History Collections.

Haywood, B. K. 2014. A "sense of place" in public participation in scientific research. *Science Education* 98(1):64–83.

Hazbón, M. H., L. Rigouts, M. Schito, M. Ezewudo, T. Kudo, T. Itoh, M. Ohkuma, K. Kiss, L. Wu, J. Ma, M. Hamada, M. Strong, M. Salfinger, C. L. Daley, J. A. Nick, J. S. Lee, N. Rastogi, D. Couvin, R. Hurtado-Ortiz, C. Bizet, A. Suresh, T. Rodwell, A. Albertini, K. A. Lacourciere, A. Deheer-Graham, S. Alexander, J. E. Russell, R. Bradford, and M. A. Riojas. 2018. Mycobacterial biomaterials and resources for researchers. *Pathogens and Disease* 76(4):fty042.

Heberling, J. M., and B. L. Isaac. 2017. Herbarium specimens as exaptations: New uses for old collections. *American Journal of Botany* 104(7):963–965.

Heberling, J. M., L. A. Prather, and S. J. Tonsor. 2019. The changing uses of herbarium data in an era of global change: An overview using automated content analysis. *BioScience* 69(10):812–822.

Hedrick, B. P., J. M. Heberling, E. K. Meineke, K. G. Turner, C. J. Grassa, D. S. Park, J. Kennedy, J. A. Clarke, J. A. Cook, D. C. Blackburn, S. V. Edwards, and C. C. Davis. 2020. Digitization and the future of natural history collections. *BioScience* 70(3):243–251.

Heidorn, P. B. 2008. Shedding light on the dark data in the long tail of science. *Library Trends* 57(2):280–299.

Hendy, A. J. W., and B. J. MacFadden. 2014. Digitizing paleontological collections for new audiences: Past practices and the potential for public participation. In *10th North American Paleontological Convention, Abstract Book,* Volume 13 of The Paleontological Society Special Publications. Pp. 127–128.

Hernandez, P. R., A. Woodcock, M. Estrada, and P. W. Schultz. 2018. Undergraduate research experiences broaden diversity in the scientific workforce. *BioScience* 68(3):204–211.

Hickey, J. J., and D. W. Anderson. 1968. Chlorinated hydrocarbons and eggshell changes in raptorial and fish-eating birds. *Science* 162(3850):271–273.

Hill, A., R. Guralnick, A. Smith, A. Sallans, R. Gillespie, M. Denslow, J. Gross, Z. Murrell, T. Conyers, P. Oboyski, J. Ball, A. Thomer, R. Prys-Jones, J. de la Torre, P. Kociolek, and L. Fortson. 2012. The Notes from Nature tool for unlocking biodiversity records from museum records through citizen science. *ZooKeys* 209:219–233.

Hiller, A. W., C. Cicero, M. J. Albe, T. L. W. Barclay, C. L. Spencer, M. S. Koo, R. C. K. Bowie, and E. A. Lace. 2017. Mutalisim in museums: A model for engaging undergraduates in biodiversity science. *PLOS Biology* 15(11):e2003318.

Ho, S.-H., and J. A. Cook. 2013. Co-evolving pedagogies. *ARID: A Journal of Desert Art, Design and Ecology*, March 17. https://aridjournal.com/co-evolving-pedagogies-szu-han-ho-and-joseph-a-cook (accessed September 6, 2020).

Hoberg, E. P., S. Kutz, J. Cook, K. Galaktionov, V. Haukisalmi, H. Henttonen, and S. Laaksonen. 2013. Parasites in terrestrial, freshwater and marine systems. In H. Meltofte (ed.), *Arctic biodiversity assessment: Status and trends in Arctic biodiversity*. Akureyi, Iceland: Conservation of Arctic Flora and Fauna. Pp. 476–505.

Hobern, D., B. Baptiste, K. Copas, R. Guralnick, A. Hahn, E. van Huis, E.-S. Kim, M. McGeoch, I. Naicker, L. Navarro, D. Noesgaard, M. Price, A. Rodrigues, D. Schigel, C. A. Sheffield, and J. Wieczorek. 2019. Connecting data and expertise: A new alliance for biodiversity knowledge. *Biodiversity Data Journal* 7:e33679.

Hoffmaster, A. R., C. C. Fitzgerald, E. Ribot, L. W. Mayer, and T. Popovic. 2002. Molecular subtyping of *Bacillus anthracis* and the 2001 bioterrorism-associated anthrax outbreak, United States. *Emerging Infectious Diseases* 8(10):1111–1116.

Hong, L., and S. E. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101:16385–16389.

Hooper-Greenhill, E. 1994. *Museums and their visitors*. New York: Routledge.

Humphreys, A. M., R. Govaerts, S. Z. Ficinski, E. N. Lughadha, and M. S. Vorontsova. 2019. Global dataset shows geography and life form predict modern plant extinction and rediscovery. *Nature Ecology and Evolution* 3:1043–1047.

Hunter, P. 2008. The paradox of model organisms. *EMBO Reports* 9(8):717–720.

Ingram, A. L., and A. R. Parker. 2008. A review of the diversity and evolution of photonic structures in butterflies, incorporating the work of John Huxley (the Natural History Museum, London from 1961 to 1990). *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1502):2465–2480.

Innis, M. A., K. B. Myambo, D. H. Gelfand, and M. A. Brow. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction–amplified DNA. *Proceedings of the National Academy of Sciences* 85(24):9436–9440.

IPBES (Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services). 2019. *Global assessment report on biodiversity and ecosystem services*. Bonn, Germany: IPBES.

IPCC (Intergovernmental Panel on Climate Change). *Special report on climate and land.* https://www.ipcc.ch/srccl (accessed August 24, 2020).

Ishino, Y., H. Shinagawa, K. Makino, M. Amemura, and A. Nakata. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology* 169(12):5429–5433.

ISO (International Organization for Standardization). 2015. ISO: 9001:2015: Quality management systems—requirements. https://www.iso.org/standard/62085.html (accessed November 30, 2020).

ISO. 2018. ISO 20387:2018: Biotechnology—biobanking—general requirements for biobanking. https://www.iso.org/standard/67888.html (accessed November 30, 2020).

IWGSC (Interagency Working Group on Scientific Collections). 2009. *Scientific collections: Mission-critical infrastructure of federal science agencies*. Washington, DC: Office of Science and Technology Policy.

Jahn, T., M. Bergmann, and F. Keil. 2012. Transdisciplinarity: Between mainstreaming and marginalization. *Ecological Economics* 79:1–10.

James, S. A., P. S. Soltis, L. Belbin, A. D. Chapman, G. Nelson, D. L. Paul, and M. Collins. 2018. Herbarium data: Global biodiversity and societal botanical needs for novel research. *Applications in Plant Sciences* 6(2):e1024.

Jamieson, K. H., D. Kahan, and D. Scheufele. 2017. *Oxford handbook of the science of science communication*. Oxford, UK: Oxford University Press.

Jarreau, P. B., N. S. Dahmen, and E. Jones. 2019. Instagram and the science museum: A missed opportunity for public engagement. *Journal of Science Communication* 18(2):A06.

Jarrett, R. L., and K. McCluskey (eds.). 2019. *The biological resources of model organisms*. Boca Raton, FL: CRC Press.

Jimenez, M. F., T. M. Laverty, S. P. Bombaci, K. Wilkins, D. E. Bennett, and L. Pejchar. 2019. Underrepresented faculty play a disproportionate role in advancing diversity and inclusion. *Nature Ecology & Evolution* 3:1030–1033.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337(6096):816–821.

Johnson, K. G., S. J. Brooks, P. B. Fenberg, A. G. Glover, K. E. James, A. M. Lister, E. Michel, M. Spencer, J. A. Todd, E. Valsami-Jones, J. R. Young, and J. R. Stewart. 2011. Climate change and biosphere response: Unlocking the collections vault. *BioScience* 61(2):147–153.

Johnson, P. T. J., R. S. Ostfelt, and F. Keesing. 2015. Frontiers in research on biodiversity and disease. *Ecology Letters* 18:1119–1133.

Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, and P. Daszak. 2008. Global trends in emerging infectious diseases. *Nature* 451(7181):990–993.

Kahl, L. J., and D. Endy. 2013. A survey of enabling technologies in synthetic biology. *Journal of Biological Engineering* 7(1):13.

Karesh, W. B., R. A. Cook, E. L. Bennett, and J. Newcomb. 2005. Wildlife trade and global disease emergence. *Emerging Infectious Diseases* 11:1000–1002.

Karim, T., R. Burkhalter, Ã. Farrell, A. Molineux, G. Nelson, J. Utrup, and S. Butts. 2016. Digitization workflows for paleontology collections. *Palaeontologia Electronica* 19.3.4T:1–14.

Kates, H., R. Folk, D. Conde, B. Ruben, C. Dervinis, R. LaFrance, M. Kirst, R. Guralnick, D. Soltis, and P. S. Soltis. 2018. Rapid workflows from specimens to sequences: Global-scale phylogenomics from collections. Presentation at Botany 2018 Colloquium. http://2018.botanyconference.org/engine/search/index.php?func=detail&aid=708 (accessed August 24, 2020).

Kates, H. R., J.-M. Ane, K. Balmont, D. Conde, M. Crook, C. Dervinis, R. P. Guralnick, T. Irving, M. Kirst, S. Knaack, L. Maia, S. Roy, R. Folk, D. E. Soltis, and P. S. Soltis. 2019. Global-scale phylogenomics of the nitrogen-fixing clade. Presentation at Plant & Animal Genome Conference XXVII. https://pag.confex.com/pag/xxvii/meetingapp.cgi/Paper/33580 (accessed August 24, 2020).

Katz, J. S., and B. R. Martin. 1997. What is research collaboration? *Research Policy* 26(1):1–18.

Kemp, C. 2015. Museums: The endangered dead. *Nature* 518(7539):292–294.

Khan, M. I., J. Shin, and J.-D. Kim. 2018. The promising future of microalgae: Current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. *Microbial Cell Factories* 17(1):36.

Knight, T. F. 2003. *Idempotent vector design for standard assembly of BioBricks*. MIT Synthetic Biology Working Group Technical Reports 2003. http://hdl.handle.net/1721.1/21168 (accessed August 24, 2020).

Konig, C., P. Weigelt, J. Schrader, A. Taylor, J. Kattge, and H. Kreft. 2019. Biodiversity data integration—the significance of data resolution and domain. *PLOS Biology* 17(3):e3000183.

Koornneef, M., and D. Meinke. 2010. The development of *Arabidopsis* as a model plant. *Plant Journal* 61(6):909–921.

Kraemer, M. U., M. E. Sinka, K. A. Duda, A. Q. Mylne, F. M. Shearer, C. M. Barker, C. G. Moore, R. G. Carvalho, G. E. Coelho. W. Van Bortel, G. Hendrickx, F. Schaffner, I. R. F. Elyazar, H.-J. Teng, O. J. Brady, J. P. Messina, D. M. Pigott, T. W. Scott, D. L. Smith, G. R. W. Wint, N. Golding, and S. I. Hay. 2015. The global distribution of the arbovirus vectors *Aedes aegypti* and *Ae. albopictus*. *Elife* 4:e08347.

Kreuzer, P., and D. Dreesmann. 2016. Museum behind the scenes—An inquiry-based learning unit with biological collections in the classroom. *Journal of Biological Education* 51(3):261–272.

Krishtalka, L., and P. S. Humphrey. 2000. Can natural history museums capture the future? *BioScience* 50(7):611–617.

Krishtalka, L., E. Dalcin, S. Ellis, J. Ganglo, T. Hosoya, M. Nakae, I. Owens, D. Paul, M. Pignal, B. Theirs, and S. Masinde. 2016. *Accelerating the discovery of biocollections data*. Copenhagen, Denmark: GBIF Secretariat.

Kutz, S. J., E. P. Hoberg, J. Nagy, L. Polley, and B. Elkin. 2004. "Emerging" parasitic infections in Arctic ungulates. *Integrative and Comparative Biology* 44(2):109–118.

Kvaskoff, M., and S. D. McKay. 2014. Scientists need leadership training. *Nature* 506:159.

Lab Manager. 2015. Professor's egg research hatches new discoveries on environmental change. *News*, December 2. https://www.labmanager.com/news/professor-s-egg-research-hatches-new-discoveries-onenvironmental-change-10908 (accessed November 18, 2020).

Lacey, E. A., T. T. Hammond, R. E. Walsh, K. C. Bell, S. V. Edwards, E. R. Ellwood, R. Guralnick, S. M. Ickert-Bond, A. R. Mast, J. E. McCormack, A. K. Monfils, P. S. Soltis, D. E. Soltis, and J. A. Cook. 2017. Climate change, collections and the classroom: Using big data to tackle big problems. *Evolution: Education and Outreach* 10(1):2.

Lambert, J. 2019a. Alaska governor halves massive funding cut to state university system. *Nature*, August 14. https://www.nature.com/articles/d41586-019-02462-2 (accessed July 26, 2020).

Lambert, J. 2019b. "No one is immune": Alaska's scientists despair over plan to shrink state universities. *Nature* 572:164–165.

Lane, J., and S. Bertuzzi. 2011. Measuring the results of science investments. *Science* 331(6018):678–680.

Lang, P. L. M., F. M. Willems, J. F. Scheepens, H. A. Burbano, and O. Bossdorf. 2019. Using herbaria to study global environmental change. *New Phytologist* 221(1):110–122.

Larkin, N. 2011. Pyrite decay: Cause and effect, prevention and cure. *Natural Sciences Collections Association News* 21:35–43.

Laurence, A. R., and V. M. Bryant. 2019. Forensic palynology and the search for geolocation: Factors for analysis and the Baby Doe case. *Forensic Science International* 302:109903.

Lave, J., and E. Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Ławniczak, Ł., M. Woźniak-Karczewska, A. P. Loibner, H. J. Heipieper, and Ł. Chrzanowski. 2020. Microbial degradation of hydrocarbons-basic principles for bioremediation: A review. *Molecules* 25(4):856.

Leather, S. R., and D. J. L. Quicke. 2009. Where would Darwin have been without taxonomy? *Journal of Biological Education* 43(2):51–52.

Leiserson, C. E., and C. McVinny. 2015. Lifelong learning: Science professors need leadership training. *Nature* 523(7560):279–281.

Lendemer, J., B. Thiers, A. K. Monfils, J. Zaspel, E. R. Ellwood, A. Bentley, K. LeVan, J. Bates, D. Jennings, D. Contreras, L. Lagomarsino, P. Mabee, L. S. Ford, R. Guralnick, R. E. Gropp, M. Revelez, N. Cobb, K. Seltmann, and M. C. Aime. 2020. The extended specimen network: A strategy to enhance US biodiversity collections, promote research and education. *BioScience* 70(1):23–30.

Lessard, B. D., A. L. Whiffin, and A. L. Wild. 2017. A guide to public engagement for entomological collections and natural history museums in the age of social media. *Annals of the Entomological Society of America* 110(5):467–479.

Li, W., Q. Song, R. Brlansky, and J. S. Hartung. 2007. Genetic diversity of citrus canker pathogens preserved in herbarium specimens. *Proceedings of the National Academy of Sciences* 104(47):18427–18431.

Lin, D., J. Crabtree, I. Dillo, R. R. Downs, R. Edmunds, D. Giaretta, M. De Giusti, H. L'Hours, W. Hugo, R. Jenkyns, V. Khodiyar, M. E. Martone, M. Mokrane, V. Navale, J. Petters, B. Sierman, D. V. Sokolova, M. Stockhause, and J. Westbrook. 2020. The trust principles for digital repositories. *Scientific Data* 7(1):144.

Lister, A. M. 2011. Natural history collections as sources of long-term datasets. *Trends in Ecology & Evolution* 26(4):153–154.

Lobell, D. B., W. Schlenker, and J. Costa-Roberts. 2011. Climate trends and global crop production since 1980. *Science* 313:616–620.

Locey, K. J., and J. T. Lennon. 2016. Scaling laws predict global microbial diversity. *Proceedings of the National Academy of Sciences* 113(21):5970–5975.

Lorieul, T., K. Pearson, E. Ellwood, H. Goëau, J.-F. Molino, P. Sweeney, J. Yost, J. Sachs, E. Mata-Montero, G. Nelson, P. Soltis, P. Bonnet, and A. Joly. 2019. Toward a large-scale and deep phenological stage annotation of herbarium specimens: Case studies from temperate, tropical, and equatorial floras. *Applications in Plant Sciences* 7:e01233.

Lutz, H., V. Tkach, and J. Weckstein. 2017. Methods for specimen-based studies of avian symbionts. In M. S. Webster (ed.), *The extended specimen: Emerging frontiers in collections-based ornithological research*. Studies in Avian Biology, no. 50. Boca Raton, FL: CRC Press. Pp. 157–183.

MacFadden, B. J. 2019. *Broader impacts of science on society.* Cambridge, UK: Cambridge University Press.

Machalaba, C., C. Romanelli, P. Stoett, S. E. Baum, T. A. Bouley, P. Daszak, and W. B. Karesh. 2015. Climate change and health: Transcending silos to find solutions. *Annals of Global Health* 81(3):445–458.

Malaney, J. L., and J. A. Cook. 2018. A perfect storm for mammalogy: Declining sample availability in a period of rapid environmental degradation. *Journal of Mammalogy* 99(4):773–788.

McCluskey, K. 2017. A review of living collections with special emphasis on sustainability and its impact on research across multiple disciplines. *Biopreservation and Biobanking* 15(1):20–30.

McCluskey, K., and M. Plamann. 2008. Perspectives on genetic resources at the fungal genetics stock center. *Fungal Genetics Reports* 55(1):15–17.

McCluskey, K., A. Wiest, and M. Plamann. 2010. The fungal genetics stock center: A repository for 50 years of fungal genetics research. *Journal of Bioscience* 35(1):119–126.

McCluskey, K., A. Alvarez, R. Bennett, D. Bokati, K. Boundy-Mills, D. Brown, C. T. Bull, M. Coffey, T. Dreaden, C. Duke, G. Dye, E. Ehmke, K. Eversole, K. Fenstermacher, D. Geiser, Jessie A. Glaeser, S. Greene, L. Gribble, M. P. Griffith, K. Hanser, R. Humber, B. W. Johnson, A. Kermode, M. Krichevsky, M. Laudon, J. Leach, J. Leslie, M. May, U. Melcher, D. Nobles, N. R. Fonseca, S. Robinson, M. Ryan, J. Scott, C. Silflow, A. Vidaver, K. M. Webb, J. E. Wertz, S. Yentsch, and S. Zehr. 2016. The U.S. Culture Collection Network lays the foundation for progress in preservation of valuable microbial resources. *Phytopathology* 106(6):532–540.

McCluskey, K., K. Boundy-Mills, G. Dye, E. Ehmke, G. F. Gunnell, H. Kiaris, M. Polihronakis Richmond, A. D. Yoder, D. R. Zeigler, S. Zehr, and E. Grotewold. 2017. The challenges faced by living stock collections in the USA. *Elife* 6:e24611.

McCluskey, K., K. Boundy-Mills, and G. Beattie. 2018. Complying with the Nagoya Protocol to the Convention on Biological Diversity. *SIM Industrial Microbiology News* 68(1):8–9.

McCormack, J. 2018. Op-ed: Think the museum fire in Brazil can't happen here? Think again. *Los Angeles Times*, September 9. https://www.latimes.com/opinion/op-ed/la-oe-mccormack-brazil-museum-fire-funding-20180909-story.html (accessed July 26, 2020).

McCreedy, D., and L. D. Dierking. 2013. *Cascading influences: Long term impacts of informal STEM experiences for girls.* Philadelphia, PA: Franklin Institute Science Museum.

McDonald, D. 2018. Scientists urge Te Papa to invest in collections research rather than strip them of staff. stuff, July 16. https://www.stuff.co.nz/science/105511593/scientests-urge-te-papa-to-invest-in-collectionsresearch-rather-than-strip-them-of-staff (accessed November 30, 2020).

McKie, R. 2017. Six Nobel prizes—What's the fascination with the fruit fly? *The Guardian*, October 7. https://www.theguardian.com/science/2017/oct/07/fruit-fly-fascination-nobel-prizes-genetics (accessed July 26, 2020).

McKinley, D. C., A. J. Miller-Rushing, H. L. Ballard, R. Bonney, H. Brown, S. C. Cook-Patton, D. M. Evans, R. A. French, J. K. Parrish, T. B. Phillips, S. F. Ryan, L. A. Shanley, J. L. Shirk, K. F. Stepenuck, J. F. Weltzin, A. Wiggins, O. D. Boyle, R. D. Briggs, S. F. Chapin, III, D. A. Hewitt, P. W. Preuss, and M. A. Soukup. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 208:15–28.

McLean, B. S., K. C. Bell, J. L. Dunnum, B. Abrahamson, J. P. Colella, E. R. Deardorff, J. A. Weber, A. K. Jones, F. Salazar-Miralles, and J. A. Cook. 2016. Natural history collections-based research: Progress, promise, and best practices. *Journal of Mammalogy* 97(1):287–297.

McNamara, M. E., V. Saranathan, E. R. Locatelli, H. Noh, D. E. G. Briggs, P. J. Orr, and H. Cao. 2014. Cryptic iridescence in a fossil weevil generated by single diamond photonic crystals. *Journal of the Royal Society Interface* 11:20140736.

Medin, D. L., and C. D. Lee. 2012. Diversity makes better science. *Observer*, April 27. https://www.psychologicalscience.org/observer/diversity-makes-better-science (accessed September 7, 2020).

Meehan, C. J., G. A. Goig, T. A. Kohl, L. Verboven, A. Dippenaar, M. Ezewudo, M. R. Farhat, J. L. Guthrie, K. Laukens, P. Miotto, B. Ofori-Anyinam, V. Dreyer, P. Supply, A. Suresh, C. Utpatel, D. van Soolingen, Y. Zhou, P. M. Ashton, D. Brites, A. M. Cabibbe, B. C. de Jong, M. de Vos, F. Menardo, S. Gagneux, Q. Gao, T. H. Heupink, Q. Liu, C. Loiseau, L. Rigouts, T. C. Rodwell, E. Tagliani, T. M. Walker, R. M. Warren, Y. Zhao, M. Zignol, M. Schito, J. Gardy, D. M. Cirillo, S. Niemann, I. Comas, and A. Van Rie. 2019. Whole genome sequencing of *Mycobacterium tuberculosis*: Current standards and open issues. *Nature Reviews Microbiology* 17(9):533–545.

Meineke, E. K., T. J. Davies, B. H. Daru, and C. C. Davis. 2018a. Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374(1763):20170386.

Meineke, E. K., C. C. Davis, and T. J. Davies. 2018b. The unrealized potential of herbaria for global change biology. *Ecological Monographs* 88(4):505–525.

Merritt, E. 2017. Future proofing museum business plans: Creating sustainable models for natural history research collections. *Museum* May/June, p. 17.

Metsger, D. A., and S. C. Byers. 1999. *Managing the modern herbarium: An interdisciplinary approach*. Chicago, IL: Society for the Preservation of Natural History Collections.

Miller, E. A., S. E. Lisin, C. M. Smith, and K. S. V. Houtan. 2020a. *Herbaria macroalgae* as a proxy for historical upwelling trends in central California. *Proceedings of the Royal Society B: Biological Sciences* 287(1929):20200732.

Miller, S. E., L. N. Barrow, S. M. Ehlman, J. A. Goodheart, S. E. Greiman, H. L. Lutz, T. M. Misiewicz, S. M. Smith, M. Tan, C. J. Thawley, J. A. Cook, and J. E. Light. 2020b. Building natural history collections for the twenty-first century and beyond. *BioScience* 70(8):674–687.

Mohr, S. E., Y. Hu, K. Kim, B. E. Housden, and N. Perrimon. 2014. Resources for functional genomics studies in *Drosophila melanogaster*. *Genetics* 197(1):1–18.

Monfils, A. K., K. E. Powers, C. J. Marshall, C. T. Martine, J. F. Smith, and L. A. Prather. 2017. Natural history collections: Teaching about biodiversity across time, space, and digital platforms. *Southeastern Naturalist* 16(sp10):47–57.

Moritz, C., J. L. Patton, C. J. Conroy, J. L. Parra, G. C. White, and S. R. Beissinger. 2008. Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science* 322(5899):261–264.

Morse, S. S., J. A. K. Mazet, M. Woolhouse, C. R. Parrish, D. Carroll, W. B. Karesh, C. Zambrana-Torrelio, W. I. Lipkin, and P. Daszak. 2012. Prediction and prevention of the next pandemic zoonosis. *The Lancet* 380:1956–1965.

Mujtaba, T., M. Lawrence, M. Oliver, and M. J. Reiss. 2018. Learning and engagement through natural history museums. *Studies in Science Education* 54(1):41–67.

Mullis, K. B. 1990. The unusual origin of the polymerase chain reaction. *Scientific American* 262(4):56–65.

Myers, S. S. 2018. Planetary health: Protecting human health on a rapidly changing planet. *The Lancet* 390:2860–2868.

Myers, S. S., L. Gaffikin, C. D. Golden, R. S. Ostfeld, K. H. Redford, T. H. Ricketts, W. R. Turner, and S. A. Osofsky. 2013. Human health impacts of ecosystem alteration. *Proceedings of the National Academy of Sciences* 110:18753–18760.

Nagy, L. G., Z. Merényi, B. Hegedüs, and B. Bálint. 2020. Novel phylogenetic methods are needed for understanding gene function in the era of mega-scale genome sequencing. *Nucleic Acids Research* 48(5):2209–2219.

NAML and OBFS (National Association of Marine Laboratories and Organization of Biological Field Stations). 2013. *Place-based research site strategic planning survey: Results summary*. http://www.obfs.org/fsml-future (accessed August 24, 2020).

NAS (National Academy of Sciences). 2005. *Systematics and the origin of species: On Ernst Mayr's 100th anniversary*. Washington, DC: The National Academies Press.

NAS, NAE, and IOM (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine). 2011. *Expanding underrepresented minority participation: America's science and technology talent at the crossroads*. Washington, DC: The National Academies Press.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2015. *Integrating discovery-based research into the undergraduate curriculum: Report of a convocation*. Washington, DC: The National Academies Press.

NASEM. 2016. *Science literacy: Concepts, contexts, and consequences*. Washington, DC: The National Academies Press.

NASEM. 2017a. *Communicating science effectively: A research agenda*. Washington, DC: The National Academies Press.

NASEM. 2017b. *Report series: Committee on Astrobiology and Planetary Science: Getting ready for the next planetary science decadal survey*. Washington, DC: The National Academies Press.

NASEM. 2017c. *Undergraduate research experiences for STEM students: Successes, challenges, and opportunities*. Washington, DC: The National Academies Press.

NASEM. 2018a. *A midterm assessment of implementation of the decadal survey on life and physical sciences research at NASA*. Washington, DC: The National Academies Press.

NASEM. 2018b. *Data science for undergraduates: Opportunities and options*. Washington, DC: The National Academies Press.

NASEM. 2018c. *Indicators for monitoring undergraduate STEM education*. Washington, DC: The National Academies Press.

NASEM. 2018d. *Learning through citizen science: Enhancing opportunities by design*. Washington, DC: The National Academies Press.

NASEM. 2019a. *Evaluating the taxonomic status of the Mexican gray wolf and the red wolf*. Washington, DC: The National Academies Press.

NASEM. 2019b. *Frontiers of materials research: A decadal survey*. Washington, DC: The National Academies Press.

NASEM. 2019c. *Minority serving institutions: America's underutilized resource for strengthening the STEM workforce*. Washington, DC: The National Academies Press.

NASEM. 2019d. *Reproducibility and replicability in science*. Washington, DC: The National Academies Press.

NASEM. 2019e. *Science and engineering for grades 6–12: Investigation and design at the center*. Washington, DC: The National Academies Press.

NASEM. 2020. *A vision for NSF Earth sciences 2020–2030: Earth in time*. Washington, DC: The National Academies Press.

NatSCA (Natural Sciences Collection Alliance). 2005. *A matter of life and death: Natural science collections: Why keep them and why fund them?* http://natsca.org/sites/default/files/publications-full/A-Matter-Of-Life-And-Death.pdf (accessed August 24, 2020).

Neeman, H., H. M. Al-Azzawi, D. Brunson, W. Burke, D. Colbry, J. T. Falgout, J. W. Ferguson, S. Gesing, J. Gyllinsky, C. S. Simmons, J. L. Simms, M. Tanash, D. Voss, J. Wells, and S. Yockel. 2018. Art. 79 in Cultivating the cyberinfrastructure workforce via an intermediate/advanced virtual residence workshop. *PEARC '19: Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)*. New York: Association for Computing Machinery.

Nekola, J. C., B. T. Hutchins, A. Schofield, B. Najev, and K. E. Perez. 2019. *Caveat consumptor notitia museo:* Let the museum data user beware. *Global Ecology and Biogeography* 28(12):1722–1734.

Nelson, G., and S. Ellis. 2018. The history and impact of digitization and digital data mobilization on biodiversity research. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374(1763):20170391.

Nelson, G., D. Paul, G. Riccardi, and A. R. Mast. 2012. Five task clusters that enable efficient and effective digitization of biological collections. *Zookeys* 209:19–45.

Nelson, G., P. Sweeney, L. E. Wallace, R. K. Rabeler, D. Allard, H. Brown, J. R. Carter, M. W. Denslow, E. R. Ellwood, C. C. Germain-Aubrey, E. Gilbert, E. Gillespie, L. R. Goertzen, B. Legler, D. B. Marchant, T. D. Marsico, A. B. Morris, Z. Murrell, M. Nazaire, C. Neefus, S. Oberreiter, D. Paul, B. R. Ruhfel, T. Sasek, J. Shaw, P. S. Soltis, K. Watson, A. Weeks, and A. R. Mast. 2015. Digitization workflows for flat sheets and packets of plants, algae, and fungi. *Applications in Plant Sciences* 3(9):apps.1500065.

Nelson, G., P. Sweeney, and E. Gilbert. 2018. Use of globally unique identifiers (GUIDs) to link herbarium specimen records to physical specimens. *Applications in Plant Sciences* 6(2):e1027.

NESCent (National Evolutionary Synthesis Center). 2010. *A strategic plan for establishing a Network Integrated Biocollections Alliance*. https://digbiocol.files.wordpress.com/2010/08/niba_brochure.pdf (accessed November 30, 2020).

Nielsen, M. W., S. Alegria. L. Borjeson, H. Etzkowtiz, H. J. Falk-Krzesinki, A. Josh, E. Leahey, L. Smith-Doerr, A. W. Woolley, and L. Schiebinger. 2017. Gender diversity leads to better science. *Proceedings of the National Academy of Sciences* 114(8):1740–1742.

Nijar, N. 2016. Fire destroys natural history museum in New Delhi. *The New York Times*, April 2. https://www.nytimes.com/2016/04/27/world/asia/museum-fire-new-delhi.html (accessed November 30, 2020).

Nobel Prize Media. 2019. *Thomas H. Morgan—facts*. https://www.nobelprize.org/prizes/medicine/1933/morgan/facts (accessed August 24, 2020).

Nowogrodzki, A. 2016a. Biological collections threatened: Hiatus in U.S. National Science Foundation funding could hamper research. *Nature* 531:561.

Nowogrodzki, A. 2016b. Biological specimen troves get a reprieve. *Nature*, May 31. https://www.nature.com/news/biological-specimen-troves-get-a-reprieve-1.19995 (accessed July 26, 2020).

NRC (National Research Council). 1991. *The U.S. National Plant Germplasm System*. Washington, DC: National Academy Press.

NRC. 1993a. *A biological survey for the nation*. Washington, DC: National Academy Press.

NRC. 1993b. *Managing global genetic resources: Agricultural crop issues and policies*. Washington, DC: National Academy Press.

NRC. 1996. *National science education standards*. Washington, DC: National Academy Press.

NRC. 2002. *Geoscience data and collections: National resources in peril*. Washington, DC: The National Academies Press.

NRC. 2005. *Thinking strategically: The appropriate use of metrics for the climate change science program*. Washington, DC: The National Academies Press.

NRC. 2009. *Learning science in informal environments: People, places, and pursuits*. Washington, DC: The National Academies Press.

NRC. 2010. *Monitoring climate change impacts: Metrics at the intersection of the human and Earth systems*. Washington, DC: The National Academies Press.

NRC. 2011. *Designing the microbial research commons: Proceedings of an international symposium*. Washington, DC: The National Academies Press.

NRC. 2012a. *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

NRC. 2012b. *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: The National Academies Press.

NRC. 2013. *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

NRC. 2014a. *Convergence: Facilitating transdisciplinary integration of life sciences, physical sciences, engineering, and beyond*. Washington, DC: The National Academies Press.

NRC. 2014b. *Enhancing the value and sustainability of field stations and marine laboratories in the 21st century*. Washington, DC: The National Academies Press.

NRC. 2014c. *Furthering America's research enterprise*. Washington, DC: The National Academies Press.

NRC. 2015a. *Enhancing the effectiveness of team science*. Washington, DC: The National Academies Press.

NRC. 2015b. *Reaching students: What research says about effective instruction in undergraduate science and engineering*. Washington, DC: The National Academies Press.

NSF (National Science Foundation). 2018. *NSF FY 2018 budget request to Congress*. https://www.nsf.gov/about/budget/fy2018/pdf/01_fy2018.pdf (accessed March 18, 2018).

Nudds, J. R., and C. W. Pettitt (eds.). 1997. *The value and valuation of natural science collections: Proceedings of the International Conference, Manchester, 1995*. Bath, UK: Geological Society Publishing House.

Odsjö, T. 2006. The environmental specimen bank, Swedish Museum of Natural History—A base for contaminant monitoring and environmental research. *Journal of Environmental Monitoring* 8(8):791–794.

OECD (Organisation for Economic Co-operation and Development). 2004. *Guidance for the operation of biological research centres (BRCs): Certification and quality criteria for BRCs.* Paris, France: OECD.

OECD. 2007. *Best practice guidelines for biological resource centres*. Paris, France: OECD Publishing.

Olsen, E. 2015. Museum specimens find new life online. *The New York Times*, November 19. https://www.nytimes.com/2015/10/20/science/putting-museums-samples-of-life-on-the-internet.html (accessed July 26, 2020).

Owens, I., and K. Johnson. 2019. One world collection: The state of the world's natural history collections. *Biodiversity Information Science and Standards* 3:e38772.

Page, A. J., C. A. Cummins, M. Hunt, V. K. Wong, S. Reuter, M. T. Holden, M. Fookes, D. Falush, J. A. Keane, and J. Parkhill. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31(22):3691–3693.

Page, R. D. M. 2008. Biodiversity informatics: The challenge of linking data and the role of shared identifiers. *Briefings in Bioinformatics* 9(5):345–354.

Papers that matter. 2020. *Nature Geoscience* 13:459.

Parmesan, C. 1996. Climate and species' range. *Nature* 382(6594):765–766.

Parsons, J. P., and C. S. Duke. 2013. Strategies for developing and innovating living stocks collections: An ESA workshop report. *Bulletin of the Ecological Society of America* 94(1):118–129.

Patton, M. Q. 2018. *Facilitating evaluation: Principles in practice*. Thousand Oaks, CA: SAGE Publications.

Pawson, E., and E. K. Teather. 2002. "Geographical expeditions": Assessing the benefits of a student-driven fieldwork method. *Journal of Geography in Higher Education* 26(3):275–289.

Pearson, K. D., G. Nelson, M. F. J. Aronson, P. Bonnet, L. Brenskelle, C. C. Davis, E. G. Denny, E. R. Ellwood, H. Goëau, J. M. Heberling, A. Joly, T. Lorieul, S. J. Mazer, E. K. Meineke, B. J. Stucky, P. Sweeney, A. E. White, and P. S. Soltis. 2020. Machine learning using digitized herbarium specimens to advance phenological research. *BioScience* 70(7):610–620.

Peattie, A. M., and R. J. Full. 2007. Phylogenetic analysis of the scaling of wet and dry biological fibrillar adhesives. *Proceedings of the National Academy of Sciences* 104(47):18595–18600.

Pecl, G. T., M. B. Araújo, J. D. Bell, J. Blanchard, T. C. Bonebrake, I.-C. Chen, T. D. Clark, R. K. Colwell, F. Danielsen, B. Evengård, L. Falconi, S. Ferrier, S. Frusher, R. A. Garcia, R. B. Griffis, A. J. Hobday, C. Janion-Scheepers, M. A. Jarzyna, S. Jennings, J. Lenoir, H. I. Linnetved, V. Y. Martin, P. C. McCormack, J. McDonald, N. J. Mitchell, T. Mustonen, J. M. Pandolfi, N. Pettorelli, E. Popova, S. A. Robinson, B. R. Scheffers, J. D. Shaw, C. J. B. Sorte, J. M. Strugnell, J. M. Sunday, M.-N. Tuanmu, A. Vergés, C. Villanueva, T. Wernberg, E. Wapstra, and S. E. Williams. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* 355(6332):eaai9214.

Pennington, D. D., G. L. Simpson, M. S. McConnell, J. M. Fair, and R. J. Baker. 2013. Transdisciplinary research, transformative learning, and transformative science. *BioScience* 63(7):564–573.

Pennisi, E. 2020. Shuttered natural history museums fight for survival amid COVID-19 "heartbreak." *Science*, May 28. https://www.sciencemag.org/news/2020/05/shuttered-natural-history-museums-fight-survival-amid-covid-19-heartbreak (accessed November 30, 2020).

Peppe, D. J., D. L. Royer, B. Cariglino, S. Y. Oliver, S. Newman, E. Leight, G. Enikolopov, M. Fernandez-Burgos, F. Herrera, J. M. Adams, E. Correa, E. D. Currano, J. M. Erickson, L. F. Hinojosa, J. W. Hoganson, A. Iglesias, C. A. Jaramillo, K. R. Johnson, G. J. Jordan, N. J. Kraft, E. C. Lovelock, C. H. Lusk, U. Niinemets, J. Peñuelas, G. Rapson, S. L. Wing, and I. J. Wright. 2011. Sensitivity of leaf size and shape to climate: Global patterns and paleoclimatic applications. *New Phytologist* 190(3):724–739.

Pew Research Center. 2018. *Women and men in STEM often at odds over workplace equity*. https://www.pewsocialtrends.org/2018/01/09/diversity-in-the-stem-workforce-varies-widely-across-jobs (accessed November 18, 2020).

Pike, J., T. Bogich, S. Elwood, D. C. Finnoff, and P. Daszak. 2014. Economic optimization of a global strategy to address the pandemic threat. *Proceedings of the National Academy of Sciences* 111(52):18519–18523.

Pogue, C. D., M. J. Monfils, D. L. Cuthrell, B. W. Heumann, and A. K. Monfils. 2016. Habitat suitability modeling of the federally endangered Poweshiek skipperling in Michigan. *Journal of Fish and Wildlife Management* 7(2):359–368.

Powers, K. E., L. A. Prather, J. A. Cook, J. Woolley, H. L. Bart, Jr., A. K. Monfils, and P. Sierwald. 2014. Revolutionizing the use of natural history collections in education. *Science Education Review* 13(2):24–33.

Pretty, J. N., A. D. Noble, D. Bossio, J. Dixon, R. E. Hine, F. W. T. de Vries, and J. I. L. Morison. 2006. Resource-conserving agriculture increases yields in developing countries. *Environmental Science and Technology* 40:1114–1119.

Primack, D., C. Imbres, R. B. Primack, A. J. Miller-Rushing, and P. Del Tredici. 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. *American Journal of Botany* 91(8):1260–1264.

Prôa, M., and A. Donini. 2019. Museums, nature, and society: The use of natural history collections for furthering public well-being, inclusion, and participation. *Theory and Practice* 2. http://articles.themuseumscholar.org/tp_vol2proadonini (accessed September 7, 2020).

Rabeler, R. 2015. *Skeletal records accompanying images: Efficiency vs. later utility*. Presentation made to the annual meeting of the Society for the Preservation of Natural History Collections. https://www.idigbio.org/content/skeletal-records-accompanying-images-efficiency-vs-later-utility (accessed August 24, 2020).

Rabeler, R. K., H. T. Svoboda, B. Thiers, L. A. Prather, J. A. Macklin, L. P. Lagomarsino, L. C. Majure, and C. J. Ferguson. 2019. Herbarium practices and ethics, III. *Systematic Botany* 44(1):7–13, 17.

Rabinowitz, P., M. Scotch, and L. Conti. 2009. Human and animal sentinels for shared health risks. *Veterinaria Italiana* 45(1):23–24.

Radeck, J., K. Kraft, J. Bartels, T. Cikovic, F. Dürr, J. Emenegger, S. Kelterborn, C. Sauer, G. Fritz, and S. Gebhard. 2013. The *Bacillus* Biobrick Box: Generation and evaluation of essential genetic building blocks for standardized work with *Bacillus subtilis*. *Journal of Biological Engineering* 7(1):29.

Ramírez-Villegas, J., C. Khoury, A. Jarvis, D. G. Debouck, and L. Guarino. 2010. A gap analysis methodology for collecting crop genepools: A case study with *Phaseolus* beans. *PLOS ONE* 5:e13497.

Ramírez-Villegas, J., C. K. Khoury, H. Achicanoy, A. Mendez, M. Diaz, C. Sosa, D. Debouck, Z. Kehel, and L. Guarino. 2020. A gap analysis modelling framework to prioritize collecting for ex situ conservation of crop landraces. *Diversity and Distributions* 26:730–742.

Ratcliffe, D. A. 1967. Decrease in eggshell weight in certain birds of prey. *Nature* 215(5097):208–210.

Raup, D. M., and J. J. Sepkoski. 1982. Mass extinctions in the marine fossil record. *Science* 215(4539):1501–1503.

Ravenscroft, J., M. Liakata, A. Clare, and D. Duma. 2017. Measuring scientific impact beyond academia: An assessment of existing impact metrics and proposed improvements. *PLOS ONE* 12(3):e0173152.

Reiser, L., T. Z. Berardini, D. Li, R. Muller, E. M. Strait, Q. Li, Y. Mezheritsky, A. Vetushko, and E. Huala. 2016. Sustainable funding for biocuration: The *Arabidopsis* Information Resource (TAIR) as a case study of a subscription-based funding model. *Database (Oxford)* 2016:baw018.

Reiss, M. J., and S. D. Tunnicliffe. 2011. Dioramas as depictions of reality and opportunities for learning in biology. *Curator: The Museum Journal* 54(4):447–459.

Riddell, E. A., K. J. Iknayan, B. O. Wolf, B. Sinervo, and S. R. Beissinger. 2019. Cooling requirements fueled the collapse of a desert bird community from climate change. *Proceedings of the National Academy of Sciences* 116(43):21609–21615.

Riojas, M. A., A. Frank, S. L. Fenn, and M. H. Hazbón. 2019. Phylogenomic comparison of *Bacillus cereus* group strains to recently identified type strains supports the species reclassification of many strains. Poster presented at American Society for Microbiology Microbe meeting: American Type Culture Collection. https://www.atcc.org/en/Documents/Learning_Center/Research/Posters/2019/Phylogenomic_Comparison_of_Bacillus_cereus_Group_Strains.aspx (accessed September 7, 2020).

Ripple, W. J., C. Wolf, T. M. Newsome, P. Barnard, and W. R. Moomaw. 2020. World scientists' warning of a climate emergency. *BioScience* 70(1):8–12.

Ristaino, J. B. 2002. Tracking historic migrations of the Irish potato famine pathogen, *Phytophthora infestans*. *Microbes and Infection* 4(13):1369–1377.

Ristaino, J. B. 2020. The importance of mycological and plant herbaria in tracking plant killers. *Frontiers in Ecology and Evolution* 7:521.

Rocha, L. A., A. Aleixo, G. Allen, F. Almeda, C. C. Baldwin, M. V. L. Barclay, J. M. Bates, A. M. Bauer, F. Benzoni, C. M. Berns, M. L. Berumen, D. C. Blackburn, S. Blum, F. Bolaños, R. C. K. Bowie, R. Britz, R. M. Brown, C. D. Cadena, K. Carpenter, L. M. Ceríaco, P. Chakrabarty, G. Chaves, J. H. Choat, K. D. Clements, B. B. Collette, A. Collins, J. Coyne, J. Cracraft, T. Daniel, M. R. de Carvalho, K. de Queiroz, F. Di Dario, R. Drewes, J. P. Dumbacher, A. Engilis, M. V. Erdmann, W. Eschmeyer, C. R. Feldman, B. L. Fisher, J. Fjeldså, P. W. Fritsch, J. Fuchs, A. Getahun, A. Gill, M. Gomon, T. Gosliner, G. R. Graves, C. E. Griswold, R. Guralnick, K. Hartel, K. M. Helgen, H. Ho, D. T. Iskandar, T. Iwamoto, Z. Jaafar, H. F. James, D. Johnson, D. Kavanaugh, N. Knowlton, E. Lacey, H. K. Larson, P. Last, J. M. Leis, H. Lessios, J. Liebherr, M. Lowman, D. L. Mahler, V. Mamonekene, K. Matsuura, G. C. Mayer, H. Mays, J. McCosker, R. W. McDiarmid, J. McGuire, M. J. Miller, R. Mooi, R. D. Mooi, C. Moritz, P. Myers, M. W. Nachman, R. A. Nussbaum, D. Ó. Foighil, L. R. Parenti, J. F. Parham, E. Paul, G. Paulay, J. Pérez-Emán, A. Pérez-Matus, S. Poe, J. Pogonoski, D. L. Rabosky, J. E. Randall, J. D. Reimer, D. R. Robertson, M.-O. Rödel, M. T. Rodrigues, P. Roopnarine, L. Rüber, M. J. Ryan, F. Sheldon, G. Shinohara, A. Short, W. B. Simison, W. F. Smith-Vaniz, V. G. Springer, M. Stiassny, J. G. Tello, C. W. Thompson, T. Trnski, P. Tucker, T. Valqui, M. Vecchione, E. Verheyen, P. C. Wainwright, T. A. Wheeler, W. T. White, K. Will, J. T. Williams, G. Williams, E. O. Wilson, K. Winker, R. Winterbottom, and C. C. Witt. 2014. Specimen collection: An essential tool. *Science* 344(6186):814–815.

Rochmyaningsih, D. 2019. Indonesia's strict new biopiracy rules could stifle international research. *Science*, July 24. https://www.sciencemag.org/news/2019/07/indonesia-s-strict-new-biopiracy-rules-could-stifle-international-research (accessed July 26, 2020).

Rodrigo, A., S. Alberts, K. Cranston, J. Kingsolver, H. Lapp, C. McClain, R. Smith, T. Vision, J. Weintraub, and B. Wiegmann. 2013. Science incubators: Synthesis centers and their role in the research ecosystem. *PLOS Biology* 11(1):e1001468.

Rodríguez Mega, E. 2020. Second Brazilian museum fire in two years reignites calls for reform. *Nature* 583(7815):175–176.

Roger, E., and S. Klistorner. 2016. Bioblitzes help science communicators engage local communities in environmental research. *Journal of Science Communication* 15(03):A06.

Rogers, N. 2016. Biologists ask NSF to reconsider plan to pause collections funding program. *Science*, March 15. https://www.sciencemag.org/news/2016/03/biologists-ask-nsf-reconsider-plan-pause-collections-funding-program (accessed July 27, 2020).

Roper, M., C. Lee, P. C. Hickey, and A. S. Gladfelter. 2015. Life as a moving fluid: Fate of cytoplasmic macromolecules in dynamic fungal syncytia. *Current Opinion in Microbiology* 26:116–122.

Rose, C. L., and A. R. de Torres (eds.). 2002. *Storage of natural history collections: Ideas and practical solutions*. Chicago, IL: Society for the Preservation of Natural History Collections.

Rouhan, G., L. J. Dorr, L. Gautier, P. Clerc, S. Muller, and M. Gaudeul. 2017. The time has come for natural history collections to claim co-authorship of research articles. *Taxon* 66(5):1014–1016.

Rowe, K. C., K. M. C. Rowe, M. W. Tingley, M. S. Koo, J. L. Patton, C. J. Conroy, J. D. Perrine, S. R. Beissinger, and C. Moritz. 2015. Spatially heterogeneous impact of climate change on small mammals of montane California. *Proceedings of the Royal Society B: Biological Sciences* 282(1799):20141857.

Rubin, G. M., and E. B. Lewis. 2000. A brief history of *Drosophila*'s contributions to genome research. *Science* 287(5461):2216–2218.

Ryan, S. J., C. J. Carlson, E. A. Mordecai, and L. R. Johnson. 2019. Global expansion and redistribution of *Aedes*-borne virus transmission risk with climate change. *PLOS Neglected Tropical Diseases* 13(3):e0007213.

Sackton, T. B., P. Grayson, A. Cloutier, Z. Hu, J. S. Liu, N. E. Wheeler, P. P. Gardner, J. A. Clarke, A. J. Baker, M. Clamp, and S. V. Edwards. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* 364(6435):74–78.

Salgado-Salazar, C., N. LeBlanc, A. Ismaiel, Y. Rivera, C. Y. Warfield, and J. A. Crouch. 2018. Genetic variation of the pathogen causing impatiens downy mildew predating and including twenty-first century epidemics on *Impatiens walleriana*. *Plant Disease* 102(12):2411–2420.

Saranathan, V., J. D. Forster, H. Noh, S. F. Liew, S. G. J. Mochrie, H. Cao, E. R. Dufresne, and R. O. Prum. 2012. Structure and optical function of amorphous photonic nanostructures from avian feather barbs: A comparative small angle x-ray scattering (SAXS) analysis of 230 bird species. *Journal of the Royal Society Interface* 9(75):2563–2580.

Saranathan, V., A. E. Seago, A. Sandy, S. Narayanan, S. G. J. Mochrie, E. R. Dufresne, H. Cao, C. O. Osuji, and R. O. Prum. 2015. Structural diversity of arthropod biophotonic nanostructures spans amphiphilic phase-space. *Nano Letters* 15(6):3735–3742.

Saupe, E. E., J. R. Hendricks, A. T. Peterson, and B. S. Lieberman. 2014. Climate change and marine molluscs of the western North Atlantic: Future prospects and perils. *Journal of Biogeography* 41:1352–1366.

Saupe, E. E., H. Qiao, J. R. Hendricks, R. W. Portell, S. J. Hunter, J. Soberón, and B. S. Lieberman. 2015. Niche breadth and geographic range size as determinants of species survival on geological time scales. *Global Ecology and Biogeography* 24(10):1159–1169.

Scheffers, B. R., L. De Meester, T. C. L. Bridge, A. A. Hoffmann, J. M. Pandolfi, R. T. Corlett, S. H. M. Butchart, P. Pearce-Kelly, K. M. Kovacs, D. Dudgeon, M. Pacifici, C. Rondinini, W. B. Foden, T. G. Martin, C. Mora, D. Bickford, and J. E. M. Watson. 2016. The broad footprint of climate change from genes to biomes to people. *Science* 354(6313):aaf7621.

Schindel, D. E., and J. A. Cook. 2018. The next generation of natural history collections. *PLOS Biology* 16(7):e2006125.

Schlaeppi, K., and D. Bulgarelli. 2015. The plant microbiome at work. *Molecular Plant–Microbe Interactions* 28(3):212–217.

Schmidly, D. J. 2005. What it means to be a naturalist and the future of natural history at American universities. *Journal of Mammalogy* 86:449–456.

Schmitt, C. J., J. A. Cook, K. R. Zamudio, and S. V. Edwards. 2018. Museum specimens of terrestrial vertebrates are sensitive indicators of environmental change in the Anthropocene. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374(1763):20170387.

Schwartz, N. D. 2019. Coronavirus recession looms, its course "unrecognizable." *The New York Times*, March 21. https://www.nytimes.com/2020/03/21/business/economy/coronavirus-recession.html (accessed November 30, 2020).

Scott, G. W., R. Goulder, P. Wheeler, L. J. Scott, M. L. Tobin, and S. Marsham. 2012. The value of fieldwork in life and environmental sciences in the context of higher education: A case study in learning about biodiversity. *Journal of Science Education and Technology* 21:11–21.

Scranton, M. A., J. T. Ostrand, F. J. Fields, and S. P. Mayfield. 2015. *Chlamydomonas* as a model for biofuels and bio-products production. *The Plant Journal* 82(3):523–531.

Seltmann, K. C., S. Lafia, D. L. Paul, S. A. James, D. Bloom, N. Rios, S. Ellis, U. Farrell, J. Utrup, M. Yost, E. Davis, R. Emery, G. Motz, J. Kimmig, V. Shirey, E. Sandall, D. Park, C. Tyrrell, R. S. Thackurdeen, M. Collins, V. O'Leary, H. Prestridge, C. Evelyn, and B. Nyberg. 2018. Georeferencing for research use (GRU): An integrated geospatial training paradigm for biocollections researchers and data providers. *Research and Ideas and Outcomes* 4:e32449.

Shaffer, H. B., R. N. Fisher, and C. Davidson. 1998. The role of natural history collections in documenting species declines. *Trends in Ecology & Evolution* 13(1):27–30.

Sharma, S., S. Ciufo, E. Starchenko, D. Darji, L. Chlumsky, I. Karsch-Mizrachi, and C. L. Schoch. 2018. The NCBI biocollections database. *Database (Oxford)* 2018:bay006.

Shelton, S. Y. 2008. Byne's "disease": How to recognize, handle and store affected shells and related collections. *National Park Service Conserve O Gram* 11:1–4. https://www.nps.gov/museum/publications/conserveogram/11-15.pdf (accessed November 30, 2020).

Shi, Q., B. Xu, X. Xu, Y. Xiao, W. Wang, and H. Wang. 2011. Diversity of social ties in scientific collaboration networks. *Physica A: Statistical Mechanics and Its Applications* 390(23):4627–4635.

Shouman, S., N. Mason, J. M. Heberling, T. Kichey, D. Closset-Kopp, A. Kobeissi, and G. Decocq. 2020. Leaf functional traits at home and abroad: A community perspective of sycamore maple invasion. *Forest Ecology and Management* 464:118061.

Shrivastava, S., V. Puri, K. A. Dilley, E. Ngouajio, J. Shifflett, L. M. Oldfield, N. B. Fedorova, L. Hu, T. Williams, A. Durbin, P. Amedeo, S. Rashid, R. S. Shabman, and B. E. Pickett. 2018. Whole genome sequencing, variant analysis, phylogenetics, and deep sequencing of Zika virus strains. *Science Reports* 8(1):15843.

Sigwart, J. D. 2018. *What species mean: A user's guide to the units of biodiversity*. Boca Raton, FL: CRC Press.

Smith, D., K. McCluskey, and E. Stackebrandt. 2014. Investment into the future of microbial resources: Culture collection funding models and BRC business plans for biological resource centers. *Springerplus* 3:81.

Soltis, P. S. 2017. Digitization of herbaria enables novel research. *American Journal of Botany* 104(9):1281–1284.

Soltis, P. S., G. Nelson, A. Zare, and E. K. Meineke. 2020. Plants meet machines: Prospects in machine learning for plant biology. *Applications in Plant Sciences* 8(6):e11371.

Soul, L. C., R. S. Barclay, A. Bolton, and S. L. Wing. 2018. Fossil atmospheres: A case study of citizen science in question-driven palaeontological research. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374(1763):20170388.

SPNHC (Society for the Preservation of Natural History Collections). 1994. Guidelines for the care of natural history collections. *Collection Forum* 10:32–40.

Stackebrandt, E., M. Schüngel, D. Martin, and D. Smith. 2015. The Microbial Resource Research Infrastructure MIRRI: Strength through coordination. *Microorganisms* 3(4):890–902.

Stern, S. 2004. *Biological resource centers: Knowledge hubs for the life sciences*. Washington, DC: The Brookings Institution Press.

Stewart, C., S. Simms, B. Plale, M. Link, D. Hancock, and G. Fox. 2010. What is Cyberinfrastructure? In *Proceedings of SIGUCCS 2010* (Norfolk, VA, October 24–27). http://dsc.soic.indiana.edu/publications/fp109a-stewart.pdf (accessed November 30, 2020).

Stoner, N. 2002. *Clean water at risk: A 30th anniversary assessment of the Bush Administration's rollback of clean water protections*. New York: Natural Resources Defense Council, Clean Water Network.

Strasser, B. J. 2008. GenBank—Natural history in the 21st century? *Science* 322(5901):537–538.

Suarez, A. V., and N. D. Tsutsui. 2004. The value of museum collections for research and society. *BioScience* 54(1):66–74.

Suhrbier, L., W. H. Kusber, O. Tschöpe, A. Güntsch, and W. G. Berendsohn. 2017. AnnoSys-implementation of a generic annotation system for schema-based data using the example of biodiversity collection data. *Database (Oxford)* 2017(1):bax018.

Sullivan, H., and C. Skelcher. 2002. *Working across boundaries. Collaboration in public services*. New York: Palgrave Macmillan.

Tanabe, S. 2006. Environmental specimen bank in Ehime University (es-BANK), Japan for global monitoring. *Journal of Environmental Monitoring* 8(8):782–790.

Tanis, B. P., L. R. G. DeSantis, and R. C. Terry. 2018. Dental microwear textures across cheek teeth in canids: Implications for dietary studies of extant and extinct canids. *Palaeogeography, Palaeoclimatology, Palaeoecology* 508:129–138.

Taubenberger, J. K., A. H. Reid, A. E. Krafft, K. E. Bijwaard, and T. G. Fanning. 1997. Initial genetic characterization of the 1918 "Spanish" influenza virus. *Science* 275(5307):1793–1796.

Taunt, H. N., L. Stoffels, and S. Purton. 2018. Green biologics: The algal chloroplast as a platform for making biopharmaceuticals. *Bioengineered* 9(1):48–54.

Tewksbury, J. J., A. K. Salomon, C. M. del Rio, D. J. Levey, J. G. T. Anderson, J. D. Bakker, K. Rowell, L. Stacey, M. J. Groom, M. E. Power, N. J. Machnicki, P. W. Dunwiddie, S. E. Hampton, S. C. Trombulak, S. G. Herman, T. A. Wheeler, and T. J. Billo. 2014. Natural history's place in science and society. *BioScience* 64(4):300–310.

Thiers, B., R. Rivas, and E. Kiernan. 2018. Using data from index herbariorum to assess threats to the world's herbaria. *Biodiversity Information Science and Standards* 2:e26440.

Thiers, B., A. K. Monfils, J. Zaspel, E. Ellwood, A. Bentley, K. Levan, J. Bates, D. Jennings, D. Contreras, and L. Lagomarsino. 2019. Extending US biodiversity collections to promote research and education. *Biodiversity Collections Network* 1-28. https://bcon.aibs.org/2019/04/04/bcon-report-extending-u-s-biodiversity-collections-to-promote-research-and-education (accessed July 26, 2020).

Tollefson, J., I. de Pater, S. Luszcz-Cook, and D. DeBoer. 2019. Neptune's latitudinal variations as viewed with ALMA. *Astronomical Journal* 157(6):251.

Troudet, J., P. Grandcolas, A. Blin, R. Vignes-Lebbe, and F. Legendre. 2017. Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports* 7(1):9132.

Tschöpe, O., J. A. Macklin, R. A. Morris, L. Suhrbier, and W. G. Berendsohn. 2013. Annotating biodiversity data via the Internet. *Taxon* 62(6):1248–1258.

Tulig, M., N. Tarnowsky, M. Bevans, A. Kirchgessner, and B. Thiers. 2012. Increasing the efficiency of digitization workflows for herbarium specimens. *Zookeys* 209:103–113.

Tunnicliffe, S. D., and A. Scheersoi. 2015. Natural history dioramas: History, construction and educational role. New York: Springer.

UN DESA (United Nations Department of Economic and Social Affairs). 2019. Growing at a slower pace, world population is expected to reach 9.7 billion in 2050 and could peak at nearly 11 billion around 2100. *News*, June 17. https://www.un.org/development/desa/en/news/population/world-population-prospects-2019.html (accessed November 30, 2020).

UNEP and ILRI (United Nations Environment Programme and International Livestock Research Institute). 2020. *Preventing the next pandemic: Zoonotic diseases and how to break the chain of transmission*. Nairobi, Kenya: UNEP.

University of Vermont. 2017. Fire at historic Torrey Hall. *UVM Today*, August 3. https://www.uvm.edu/uvmnews/news/fire-historic-torrey-hall (accessed August 24, 2020).

Valantine, H. A., and F. S. Collins. 2015. National Institutes of Health addresses the science of diversity. *Proceedings of the National Academy of Sciences* 112:12240–12242.

Valin, H., R. D. Sands, D. van der Mensbrugghe, G. C. Nelson, H. Ahammad, E. Blanc, B. Bodirsky, S. Fujimori, T. Hasegawa, P. Havlik, E. Heyhoe, P. Kyle, D. Mason-D'Croz, S. Paltsev, S. Rolinski, A. Tabeau, H. van Meijl, M. von Lampe, and D. Willenbockel. 2014. The future of food demand: Understanding differences in global economic models. *Agricultural Economics* 45(1):51–67.

van Rossum, J. 2017. Blockchain for research. *Digital Science*, November 27. https://digitalscience.figshare.com/articles/Blockchain_for_Research/5607778 (accessed September 7, 2020).

Vavitsas, K., M. Fabris, and C. E. Vickers. 2018. Terpenoid metabolic engineering in photosynthetic microorganisms. *Genes* 9(11):520.

Vermeulen, S. J., D. Dinesh, S. M. Howden, L. Cramer, and P. K. Thornton. 2018. Transformation in practice: A review of empirical cases of transformational adaptation in agriculture under climate change. *Frontiers in Sustainable Food Systems* 2:65.

Verslyppe, B., W. De Smet, B. De Baets, P. De Vos, and P. Dawyndt. 2014. StrainInfo introduces electronic passports for microorganisms. *Systematic and Applied Microbiology* 37(1):42–50.

Vicente-Sáez, R., and C. Martínez-Fuentes. 2018. Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research* 88:428–436.

Vincent, H., A. Amri, N. P. Castañeda-Álvarez, H. Dempewolf, E. Dulloo, L. Guarino, D. Hole, C. Mba, A. Toledo, and N. Maxted. 2019. Modeling of crop wild relative species identifies areas globally for in situ conservation. *Communications Biology* 2:136.

Visaggi, C. C. 2020. *Equity, culture, and place in teaching paleontology: Student-centered pedagogy for broadening participation*. Cambridge, MA: Cambridge University Press.

Vo, A. T. E., M. S. Bank, J. P. Shine, and S. V. Edwards. 2011. Temporal increase in organic mercury in an endangered pelagic seabird assessed by century-old museum specimens. *Proceedings of the National Academy of Sciences* 108(18):7466–7471.

Vollmar, A., Formerly, J. Macklin, and L. Ford. 2010. Natural history specimen digitization: Challenges and concerns. *Biodiversity Informatics* 7:93–112.

Vukusic, P., and J. R. Sambles. 2003. Photonic structures in biology. *Nature* 424(6950):852–855.

Wade, W., H. Thompson, A. Rybalka, and S. Vartoukian. 2016. Uncultured members of the oral microbiome. *Journal of the California Dental Association* 44(7):447–456.

Wallace, M. C., and H. J. Curzer. 2013. Moral problems and perspectives for ecological field research. *ILAR Journal* 54(1):3–4.

Wandeler, P., P. E. Hoeck, and L. F. Keller. 2007. Back to the future: Museum specimens in population genetics. *Trends in Ecology & Evolution* 22(12):634–642.

Wang, B., J. Wang, W. Zhang, and D. R. Meldrum. 2012. Application of synthetic biology in cyanobacteria and algae. *Frontiers in Microbiology* 3:344.

Wang, Y., and T. G. Lilburn. 2009. Biological resource centers and systems biology. *BioScience* 59(2):113–125.

Wangler, M. F., S. Yamamoto, and H. J. Bellen. 2015. Fruit flies in biomedical research. *Genetics* 199(3):639–653.

Warschefsky, E., R. V. Penmetsa, D. R. Cook, and E. J. von Wettberg. 2014. Back to the wilds: Tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *American Journal of Botany* 101(10):1791–1800.

Watanabe, M. E. 2017. The Nagoya Protocol: Big steps, new problems. *BioScience* 67(4):400.

Webster, M. S. (ed.). 2017. *The extended specimen: Emerging frontiers in collections-based ornithological research*. Boca Raton, FL: CRC Press.

Wenger, E. 1998. *Communities of practice: Learning, meaning, and identity*. New York: Cambridge University Press.

Wenger, E. 2000. *Communities of practice and social learning system*s. Organization 7(2):225–246.

Wenger, E., R. A. McDermott, and W. Snyder. 2002. *Cultivating communities of practice: A guide to managing knowledge*. Boston, MA: Harvard Business School Press.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson, and D. Vieglais. 2012. Darwin Core: An evolving community-developed biodiversity data standard. *PLOS ONE* 7(1):e29715.

Wiest, A., M. Plamann, and K. McCluskey. 2008. Demonstration that the *Neurospora crassa* mutation *un-4* is a single nucleotide change in the *tim16* gene encoding a subunit of the the mitochondrial inner membrane translocase. *Fungal Genetics Reports* 55:37–39.

Wildt, D. E. 2000. Genome resource banking for wildlife research, management, and conservation. *ILAR Journal* 41(4):228–234.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B. Mons. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3:160018.

Willemse, L., E. van Egmond, V. Runnel, H. Saarenmaa, A. C. Rubio, K. Gödderz, and X. Vermeersch. 2019. Future challenges in digitisation of private natural history collections. *Biodiversity Information Science and Standards* 3:e37640.

Willis, C. G., E. R. Ellwood, R. B. Primack, C. C. Davis, K. D. Pearson, A. S. Gallinat, J. M. Yost, G. Nelson, S. J. Mazer, N. L. Rossington, T. H. Sparks, and P. S. Soltis. 2017. Old plants, new tricks: Phenological research using herbarium specimens. *Trends in Ecology & Evolution* 32(7):531–546.

Wilson, E. O. 1998. *Consilience: The unity of knowledge*. New York: Knopf.

Winker, K. 2008. What I do: Notes from the frontiers of academic curating in biology. *Curator: The Museum Journal* 51(4):393–406.

Winker, K., and J. J. Withrow. 2013. Small collections make a big impact. *Nature* 493(7433):480.

Woodward, F. I. 1987. Stomatal numbers are sensitive to increases in CO_2 from pre-industrial levels. *Nature* 327(6123):617–618.

Wu, L., Q. Sun, H. Sugawara, S. Yang, Y. Zhou, K. McCluskey, A. Vasilenko, K.-I. Suzuki, M. Ohkuma, Y. Lee, V. Robert, S. Ingsriswang, F. Guissart, P. Desmeth, and J. Ma. 2013. Global Catalogue of Microorganisms (GCM): A comprehensive database and information retrieval, analysis, and visualization system for microbial resources. *BMC Genomics* 14:933.

Wu, L., Q. Sun, P. Desmeth, H. Sugawara, Z. Xu, K. McCluskey, D. Smith, V. Alexander, N. Lima, M. Ohkuma, V. Robert, Y. Zhou, J. Li, G. Fan, S. Ingsriswang, S. Ozerskaya, and J. Ma. 2017. World Data Centre for Microorganisms: An information infrastructure to explore and utilize preserved microbial strains worldwide. *Nucleic Acids Research* 45(D1):D611–D618.

Wu, L., L. Liu, Q. Sun, B. Liu, J. Yu, P. Xu, F. Chen, Y. Chen, Z. Li, H. Shi, Y. Zhou, and J. Ma. 2018. The status and future aspects of the researches of microbial resources in China. *Acta Microbiologica Sinica* 58(12):2123–2133.

Yanagihara, R., S. H. Gu, S. Arai, H. J. Kang, and J. W. Song. 2014. Hantaviruses: Rediscovery and new beginnings. *Virus Research* 187:6–14.

Yates, T. L., J. N. Mills, C. A. Parmenter, T. G. Ksiazek, R. R. Parmenter, J. R. Vande Castle, C. H. Calisher, S. T. Nichol, K. D. Abbott, J. C. Young, M. L. Morrison, B. J. Beaty, J. L. Dunnum, R. J. Baker, J. Salazar-Bravo, and C. J. Peters. 2002. The ecology and evolutionary history of an emergent disease: Hantavirus pulmonary syndrome. *BioScience* 52(11):989–998.

Zeikus, J. G., and T. D. Brock. 1971. Protein synthesis at high temperatures: Aminoacylation of tRNA. *Biochimica et Biophysica Acta—Nucleic Acids and Protein Synthesis* 228(3):736–775.

Zhang, W., J. Gu, Q. Liu, H. Su, T. Fan, and D. Zhang. 2014. Butterfly effects: Novel functional materials inspired from the wings scales. *Physical Chemistry Chemical Physics* 16(37):19767–19780.

Appendix A
Statement of Task

An ad hoc committee of the National Academies of Sciences, Engineering, and Medicine will review the role of biological collections in research and education that are supported by the National Science Foundation and develop a set of options for their future maintenance to enable their continued use to benefit science and society. For this task, biological collections are defined as living stocks (organisms) and preserved repositories of biodiversity specimens and materials. The committee will review the past and present contributions of biological collections to research and education, describe the major advances in their use over the past decade, and envision future innovative ways in which biological collections can be utilized to further advance science over the next decade. The committee will also describe the greatest challenges to maintaining biological collections and suggest a range of long-term strategies that could be used for their sustained support, individually or in groups, of research and education. In particular, the committee will:

1. Examine the past and present contributions of biological collections of all sizes and across institution types to research and education, including new types of collections and research resources that users have derived through new technologies.
2. Describe how the quality, format, and accessibility of digitized data impact the use of biological collections for research and education. Examine whether the investments by the National Science Foundation and other U.S. federal agencies in digital data and metadata have been integrated with common standards that support increased accessibility, and recommend strategies to achieve such integration.
3. Describe potential future innovative applications of biological collections to advance research over the next decade, and outline strategies to facilitate the use of collections to open new avenues of inquiry and address issues of broad societal importance, such as global environmental change, food security, conservation, and the bioeconomy.
4. Highlight how project-based collections resulting from individual research funded projects might be identified and preserved. Address challenges of how project-based collections (i.e., those maintained by individual researchers or labs) are accessioned into archival collections maintained by institutions as a generation of active researchers reach retirement.

5. Outline critical challenges to and needs for the use and maintenance of biological collections for research and education, including:
   a. scientific and technical capabilities
   b. tools and technologies
   c. facilities (e.g., space)
   d. personnel with required expertise
   e. sustainable financial resources
6. Describe the quality control challenges for living stock collections of microbes, vertebrates, model plants (e.g., *Arabidopsis*), etc., for which consistent genetic identity is crucial for research, and consider how these challenges could be addressed.
7. Examine current efforts to sustain biodiversity and living stocks collections, from small and specialized to large and endowed collections, and recommend a range of options for how to address the issue of financial sustainability.
8. Describe best practices and metrics that will enable institutions with biological collections to monitor, assess, and modify the value and impact of their collections and their strategies to facilitate their continued use for research and education.

The committee will produce a consensus study report addressing these points.

Appendix B
Public Meeting Agendas

These in-person public meetings held by the committee served as information-gathering sessions. They are listed in chronological order. The locations of in-person meetings are provided. Presentations that were made via the Internet at the in-person public meetings are noted.

MEETING 1
National Academy of Sciences—Keck Center, Room 209
500 Fifth Street, NW
Washington, DC
December 6–7, 2018

DECEMBER 6, 2018

3:15 Welcome and Introductions—Jim Collins and Shirley Pomponi, Committee Co-Chairs

3:30 Sponsors' Perspectives on the Context and Expectations for the Study—Muriel Poston and Roland Roberts, National Science Foundation

4:30 Public Comments—Members of the public are invited to share evidence and views they would like for the committee to take into consideration. Advance sign-up is required.

4:45 Adjourn Open Session

DECEMBER 7, 2018

8:15 Welcome and Introductions—Jim Collins and Shirley Pomponi, Committee Co-Chairs

8:25 Broad Considerations for the Study Outcomes

Futureproofing Natural History Collections—Elizabeth Merritt, Vice President of Strategic Foresight and Founding Director, Center for the Future of Museums

	Perspective on Sustaining Living Microbial Germplasm Repositories—Kevin McCluskey, Research Professor and Curator, Fungal Genetics Stock Center, Kansas State University (by videoconference)
9:15	Adjourn Open Session

MEETING 2
National Academy of Sciences, Room 120
2101 Constitution Avenue, NW
Washington, DC
February 7–8, 2019

FEBRUARY 7, 2019

1:30	Welcome and Introductions—Jim Collins and Shirley Pomponi, Committee Co-Chairs
1:35	History of Natural History Collections in the United States—Pamela M. Henson, Smithsonian Institution
2:00	An Overview on the Interagency Working Group on Scientific Collections—Scott E. Miller, Smithsonian Institution
2:25	Key Components of Sustainable Mission and Infrastructure for a Biological Collection—Sarah B. George, Natural History Museum of Utah
2:50	Panel Discussion with Dr. Henson, Dr. Miller, and Dr. George
3:10	Break
3:25	Leveraging Collections to Advance Science, Technology, Engineering, and Mathematics Education—Jay Labov, Senior Advisor for Education and Communication, National Academies of Sciences, Engineering, and Medicine (Retired)
3:45	Question and Answer Session with Jay Labov
3:55	Biological Collections for Understanding Biodiversity in the Anthropocene—Emily K. Meineke, Harvard University Herbaria & Libraries

4:15 Leveraging Collections to Assess Global Status of Pollinators—Ignasi Bartomeus, Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas (by videoconference)

4:35 Panel Discussion with Dr. Meineke and Dr. Bartomeus

4:50 Public Comments—Members of the public are invited to share evidence and views they would like for the committee to take into consideration. Advance sign-up is required.

5:00 Adjourn Open Session

FEBRUARY 8, 2019

8:30 Welcome and Opening Remarks

8:35 Global Catalogue of Microorganisms (GCM): The Global Cooperation Network for Culture Collections Worldwide—Juncai Ma, Chinese Academy of Sciences (by videoconference)

8:55 Question and Answer Session with Juncai Ma

9:10 The Effect of the Nagoya Protocol on Biological Collections—Breda M. Zimkus, Harvard University

9:30 Panel Discussion with Breda M. Zimkus and Dr. Ma

9:45 Public Comments—Members of the public are invited to share evidence and views they would like for the committee to take into consideration. Advance sign-up is required.

9:50 Adjourn Open Session

MEETING 3
Arnold and Mabel Beckman Center,
Huntington and Board Rooms
100 Academy Way
Irvine, CA 92617
April 23–24, 2019

APRIL 23, 2019

9:00 Updates on Federal and National Efforts: Biocollections and Biosecurity—Diane DiEuliis, Senior Research Fellow, National Defense University (by videoconference)

10:00 Committee Discussion

4:15 Welcome and Introductions—Jim Collins and Shirley Pomponi, Committee Co-Chairs

4:20 Data Integration and Attribution—Donald Hobern, Executive Secretary, International Barcode for Life Consortium (by videoconference)

5:00 Adjourn Open Session

APRIL 24, 2019

9:00 Welcome and Introductions—Jim Collins and Shirley Pomponi, Committee Co-Chairs

9:10 Perspective on the Biodiversity Literacy in Undergraduate Education Data Initiative and the Contribution of Small Collections—Anna Monfils, Director, Central Michigan University Herbarium

9:30 Arthropod Holdings and Digitization Efforts for North America with a Focus on the United States: Meeting National to Global Needs for Biodiversity Data—Neil Cobb, Director, Merriam-Powell Center for Environmental Research, Northern Arizona University (by videoconference)

9:50 Question and Answer Session with Dr. Monfils and Dr. Cobb

10:15 Long-Term Success and Challenges in Establishing and Sustaining University Museum Biological Collections—Michael Nachman, Director, Museum of Vertebrate Zoology, University of California, Berkeley (by videoconference)

10:35 Question and Answer Session with Dr. Nachman

10:50 Long-Term Success and Challenges in Establishing and Sustaining a Botanical Garden and a Seed Bank Promoting Research, Conservation, and Education—Lucinda McDade, Director of Research, California Botanic Garden

11:10 Agricultural Genebanks: Management, Use, and Challenges—Stephanie Greene, Supervisory Plant Physiologist, Department of Agriculture National Laboratory for Genetics Resources Preservation, Fort Collins (by videoconference)

11:30 Question and Answer Session with Dr. McDade and Dr. Greene

11:45 Public Comments—Members of the public are invited to share evidence and views they would like for the committee to take into consideration. Advance sign-up is required.

12:00 Adjourn Open Session. Lunch with speakers.

Appendix C
Webinars

Requests for public access to webinar presentations and written materials submitted to the committee may be submitted through the National Academies Projects and Activities Repository.

WEBINARS

1. A Philosophical Perspective on Biological Collections (February 15, 2019)
   - Rachel A. Ankeny, The University of Adelaide, Australia
   - Sabina Leonelli, University of Exeter, United Kingdom

2. Exploring the Application of Blockchain to Natural History Collections Data (May 16, 2019)
   - Nelson Rios, Yale University

3. CSIRO's National Biological Collections as 21st Century Research Infrastructure (May 24, 2019)
   - Andrew Young, National Research Collections Australia

4. Opportunities and Challenges to Expanding Access to Collections: Cultural and Legal Perspectives (July 3, 2019)
   - Todd Kuiken, North Carolina State University—"Broad Perspectives on the Access and Benefit-Sharing and Propertization of Genetic Resources"
   - Margo Bagley, Emory University School of Law—"The Nagoya Protocol and Digital Sequence Information (DSI) on Genetic Resources: Emerging Issues"
   - Christina Agapakis, Ginkgo Bioworks—"Exploring Extinct Biodiversity: Using Synthetic Biology to Revive a Lost Scent"

5. The Costs and Value of Federal Scientific Collections (July 9, 2019)
   - Keith Crane, Science and Technology Policy Institute
   - Lauren Bartels, Science and Technology Policy Institute
   - Thomas Olszewski, Science and Technology Policy Institute

Appendix D
Biographical Sketches of Committee Members and Staff

From left to right: First row: Shirley A. Pomponi (Co-Chair), Joseph A. Cook, Pamela S. Soltis, Jessica De Mouy (Staff), Barbara M. Thiers; **Second row:** Talia S. Karim, Lynn D. Dierking, Kyria Boundy-Mills; **Third row:** Audrey Thévenon (Study Director), Rick E. Borchelt, Keegan Sawyer (Staff), George I. Matsumoto; **Fourth row:** James P. Collins (Co-Chair), Scott V. Edwards, Andrew C. Bentley, Manzour H. Hazbón

James P. Collins (*Co-Chair*) is the Virginia M. Ullman Professor of Natural History and the Environment in the School of Life Sciences at Arizona State University (ASU). He is an evolutionary ecologist whose research group studies host–pathogen biology and its relationship to the decline of species, at times even to extinction. The intellectual and institutional factors that have shaped ecology's development as a science as well as ecological ethics are other research foci. From 1989 to 2002, he was the chairman of ASU's Zoology, then Biology, Department. At the National Science Foundation (NSF), Dr. Collins was the director of the Population Biology and Physiological Ecology program from 1985 to 1986. He joined NSF's senior management in 2005 serving as the assistant director for biological sciences from 2005 to 2009. Within the NSF Directorate for Biological Sciences, he oversaw a research and education portfolio that spanned molecular and cellular biosciences to global change as well as biological infrastructure. Dr.

Collins currently serves as the chair of the Board on Life Sciences of the National Academies of Sciences, Engineering, and Medicine.

Shirley A. Pomponi (*Co-Chair*) is a research professor at the Florida Atlantic University Harbor Branch Oceanographic Institute and a professor of marine biotechnology in the Bioprocess Engineering Group at Wageningen University & Research, The Netherlands. Dr. Pomponi received her Ph.D. in biological oceanography from the University of Miami. Her research focuses on marine biotechnology, and in particular, the development of sponge cell models to study how and why sponges produce chemicals with pharmaceutical relevance. She served on the President's Panel on Ocean Exploration, was the vice chair of the National Research Council's (NRC's) Committee on Exploration of the Seas, and co-chaired the NRC's consensus study report on ocean science priorities for the next decade, *Sea Change: 2015–2025 Decadal Survey of Ocean Sciences*. She is also a member of the National Science Foundation Advisory Committee for Geosciences.

Andrew C. Bentley is a collection manager of ichthyology as well as the bioinformatics manager for the Biodiversity Institute & Natural History Museum at The University of Kansas and the usability lead for the Specify collections management software project. He has an interest in marine fish as well as all things collections (primarily alcohol preserved and cryogenic tissue collections) and databases. His research interests include collection management, specifically of preservation, digitization, databasing, and maintenance of wet and cryogenic collections. Mr. Bentley also has an interest in database development and usability. Mr. Bentley also served as the president of the Society for the Preservation of Natural History Collections and is a member of the American Society of Ichthyologists and Herpetologists. He is also affiliate faculty of The University of Kansas Museum Studies program. He earned his M.Sc. in zoology from the University of Port Elizabeth, South Africa, in 1996. He has not previously served on a National Academies of Sciences, Engineering, and Medicine committee.

Rick E. Borchelt is the director of communications and public affairs for the Department of Energy's (DOE's) Office of Science, which represents a $6.5 billion portfolio supporting the basic physical sciences. In addition to DOE, his career in science, communications, and public policy includes stints at five other federal science agencies (Department of Agriculture, National Institutes of Health, National Aeronautics and Space Administration, U.S. Information Agency, and Smithsonian Institution [where he was a graduate student curatorial assistant in the Lepidoptera collection]) and tours of duty as a congressional committee press secretary and as special assistant for public affairs in the Executive Office of the President/Office of Science and Technology Policy. His experience also reflects work for the National Academy of Sciences, Johns Hopkins University, Vanderbilt University, Massachusetts Institute of Technology/the Whitehead Institute for Biomedical Research, and the University of Maryland. He was a member of the National Academies of Sciences, Engineering, and Medicine's Roundtable on Public Interfaces in the Life Sciences, and served on the National Academy of Engineering's study of engineering communication. He currently serves on the editorial board of the peer-reviewed journal *Science Communication*. He is a contract instructor for Graduate School USA in the Natural History Field Studies certificate program, jointly managed by the Audubon Naturalist Society of the Central Atlantic States. Areas of particular interest include trust in science, extension communication research, natural history citizen science, adult science learning in informal settings, and developing community-based public engagement in science.

Kyria Boundy-Mills is a curator of Phaff Yeast Culture Collection, Food Science and Technology, at the University of California, Davis. Dr. Boundy-Mills's professional expertise involves the study and expansion of the use of the Phaff collection. She has utilized and expanded the biodiversity of the Phaff collection to expand knowledge of interactions of agricultural insect pests with yeast, oleaginous (high lipid) yeasts, tolerance of yeasts to stresses including ionic liquids, and food fermentations. These publications each used numerous yeast strains, one using 180 strains belonging to more than 100 different species. Since 2013, Dr. Boundy-Mills has served on the executive board of the World Federation for Culture Collections (WFCC). Responsibilities include screening and approving new WFCC member collections, convening international conferences, and developing international standards for culture collection management. Since 2011, she has been on the steering committee of the National Science Foundation–funded United States Culture Collection Network (USCCN) led by Kevin McCluskey, Fungal Genetics Stock Center curator (Kansas State University). USCCN coordinates and promotes microbial culture collections in the United States. She hosted the fall 2014 USCCN meeting at the University of California, Davis. Through these avenues, she has learned of and promoted awareness of emerging issues affecting microbial culture collections and their users, especially the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity (the Nagoya Protocol). She has co-authored numerous publications alerting the scientific public, especially U.S. microbiologists, about Nagoya Protocol legislation. Dr. Boundy-Mills earned her Ph.D. in biochemistry from the University of Minnesota, Minneapolis, in 1992. She has not previously served on a National Academies of Sciences, Engineering, and Medicine committee.

Joseph A. Cook is the Regents Professor of Biology and the curator of mammals at the Museum of Southwestern Biology at The University of New Mexico (UNM). Previously, he held tenured faculty and curatorial positions at the University of Alaska Fairbanks (1990–2001), was the chair of biology at Idaho State University (2000–2003), and the director of the Museum of Southwestern Biology (2011–2017). Dr. Cook's research is highly collaborative and focuses on conservation, molecular evolution, and systematics, producing more than 190 peer-reviewed publications, including *Recent Mammals of Alaska* (University of Alaska Press, 2010). He held the Fulbright Fellowship in Uruguay (1993), Rotary Fellowship in Bolivia (1997), and Sitka Sound Science Center Fellowship (2013); was awarded the American Society of Mammalogists' Joseph Grinnell Award in 2016; and was appointed UNM Regents Professor in 2018. He was the president of the Natural Science Collections Alliance (2016–2017) and chaired the international AIM-UP! Research Coordination Network, which explored new ways to integrate collections-based digital resources into education initiatives. Moving from the 10th largest mammal collection in the United States when he assumed leadership in 2003, the Division of Mammals has nearly tripled in size and is now ranked third in size worldwide. Until 2017, he was also the curator of genomic resources, a frozen tissue collection for mammals that is unrivaled worldwide for size, diversity, global coverage, or the number of peer-reviewed papers on genomes, viruses and other topics that it produces annually (ca. 70). Over 25 years, he led two international field projects, one that sampled mammals and their parasites across more than 250 remote sites in Alaska, Canada, Mongolia, and Siberia and aimed to understand the biogeography of Beringia (the Beringian Coevolution Project). The other effort, the Island Surveys to Locate Endemic Species project, focused on the mammals and parasites of the Alexander Archipelago of Southeast Alaska, including the incomparable Tongass National Forest. Dr. Cook received his Ph.D. from UNM.

Lynn D. Dierking is a Sea Grant Professor in Free-Choice/Informal STEM Learning in the Colleges of Science and Education at Oregon State University, and the director of strategy and partnerships at the Institute for Learning Innovation. Her research focuses on lifelong learning, particularly free-choice learning (in afterschool, home-, and community-based contexts, such as museums and libraries), with an emphasis on youth and families, particularly those living in poverty, and/or not historically engaged in science, technology, engineering, and mathematics (STEM) learning across their lifetime. Dr. Dierking publishes extensively and is on the editorial boards for *Connected Science Learning*, *Afterschool Matters*, and the *Journal of Museum Management and Curatorship*. Dr. Dierking received a Ph.D. in science education in 1987 from the University of Florida. She received the 2016 Distinguished Contributions to Science Education through Research Award from NARST, an international organization supporting research on science learning and teaching, recognizing her contributions to, and creation of, a research field focused on lifelong, free-choice/informal learning. Dr. Dierking was a 2013 Education & Human Resources Distinguished Lecturer at the National Science Foundation, in recognition of her leadership within the STEM education field. She was also a 2011 Department of State Distinguished Keynote Speaker for International Council on Museums meetings in Brno, Czech Republic, and the U.S. Embassy in Prague. She received a 2010 John Cotton Dana Award for Leadership from the American Alliance of Museums, the highest honor bestowed to a person outside the museum field, who exhibits outstanding leadership and promotes the educational responsibility and capacity of museums. She was also on the 2006 Centennial Honor Roll of the American Alliance of Museums as one of 100 leaders who had provided leadership and service to the field throughout their careers.

Scott V. Edwards (NAS) is the Alexander Agassiz Professor of Zoology and the curator of ornithology in the Museum of Comparative Zoology at Harvard University. He joined Harvard in December 2003 after serving as faculty for 9 years in the Zoology Department and the Burke Museum at the University of Washington at Seattle. His research focuses on diverse aspects of avian biology, including evolutionary history and biogeography, disease ecology, population genetics, and comparative genomics. He has conducted fieldwork in phylogeography in Australia since 1987 and conducted some of the first phylogeographic analyses based on DNA sequencing. He did a postdoctoral fellowship in immunogenetics at the University of Florida and gained experience with studying the major histocompatibility complex (MHC) of birds, an important gene complex for interactions of birds and infectious diseases, pathogens, and mate choice. His work on the MHC led him to study the large-scale structure of the avian genome and informed his current interest in using comparative genomics to study the genetic basis of phenotypic innovation in birds. In the past 10 years, Dr. Edwards has helped develop novel methods for estimating phylogenetic trees from multilocus DNA sequence data. His recent work uses comparative genomics in diverse contexts to study macroevolutionary patterns in birds, including the origin of feathers and the evolution of flightlessness. From 2013 to 2015 Dr. Edwards served as the director of the Division of Biological Infrastructure at the National Science Foundation (NSF), overseeing funding programs focused on undergraduate research, postdoctoral fellowships, natural history collections and field stations, and cyber- and other infrastructure for all areas of biology. He served as the president of three international scientific societies based in the United States—the Society for the Study of Evolution, the Society of Systematic Biologists, and the American Genetic Association—each of which publishes a scientific journal and has memberships ranging from 500 to 2,500 scientists and students. He has served on the National Geographic Society's Committee for Research and Exploration, the Senior Advisory Boards of the NSF-funded U.S. National Evolutionary Synthesis Center and the National Institute for Mathematical and Biological Synthesis, and on the Advisory Boards of the National Museum of Natural History at the Smithsonian Institution and the Cornell Lab of Ornithology. He oversees a program funded by NSF to increase the diversity of undergraduates in evolutionary biology and biodiversity science. He is a member of the American Academy of Arts & Sciences (2009), a fellow of the American Association for the Advancement of Science (2009), and a member of the National Academy of Sciences (2015). Dr. Edwards currently serves as a member of the National Academies of Sciences, Engineering, and Medicine's Board on Life Sciences.

Manzour H. Hazbón is a senior scientist at the American Type Culture Collection (ATCC), overseeing ATCC's bacteriology laboratory operations and bioresources since 2013. Through his leadership position at ATCC, Dr. Hazbón employs a combination of microbiological knowledge and modern laboratory techniques to support infectious disease research. Dr. Hazbón represents ATCC in several national and international scientific meetings through presentations of his scientific research findings and as a subject-matter expert for ATCC in global culture collection meetings. Dr. Hazbón is actively participating with the World Federation for Culture Collections, the United States Culture Collection Network, and the World Catalogue for Microorganisms. Dr. Hazbón has devoted most of his professional career to developing molecular assays to detect and identify respiratory pathogens, and in the study of the molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. Prior to ATCC, Dr. Hazbón was a senior scientist for Meso Scale Diagnostics, LLC. In addition, Dr. Hazbón served as a microbial genomes curator for the National Institutes of Health from 2006 to 2008 and as a senior diagnostic laboratory scientist with the Walter Reed Army Institute of Research from 2008 to 2010. Dr. Hazbón received both his Ph.D. and M.Sc. in molecular biology from the Free University of Brussels and his B.Sc. in microbiology from the Universidad de los Andes.

Talia S. Karim is the collection manager for invertebrate paleontology and paleobotany at the University of Colorado Boulder Museum of Natural History (2010 to present) and was previously the invertebrate paleontology collection manager at The University of Kansas Biodiversity Institute & Natural History Museum (2008–2010). Her research interests include trilobite systematics, biostratigraphy, taphonomy, museum collections care and management, digitization of collections, and cyberinfrastructure as related to sharing museum data. Dr. Karim's interest in collections management extends into the classroom and she has taught, or co-taught, collections management–related courses for the museum studies programs at the University of Colorado and The University of Kansas. She is an active Society for the Preservation of Natural History Collections member and is currently serving as a member-at-large. She is also the co-chair of the Integrated Digitized Biocollections Paleo Digitization Working Group. Dr. Karim received a B.S. in geology and a B.A. in classical culture from The University of Oklahoma in 2001. She went on to attend Oxford University on a Marshall Scholarship and earned an M.Sc. in Earth science in 2004. She completed her Ph.D. at The University of Iowa in 2009 focusing on Lower Ordovician trilobite systematics. Throughout her career, she has been a specimen-based researcher and focused on the critical role specimens and museum collections play in research and communicating science to the general public.

George I. Matsumoto is currently the senior education and research specialist at the Monterey Bay Aquarium Research Institute (MBARI). With an A.B. from the University of California, Berkeley, and a Ph.D. from the University of California, Los Angeles, Dr. Matsumoto's research interest focuses on ctenophores but includes other gelatinous organisms, especially those that live in the deep sea. He also coordinates the MBARI summer internship program and educator professional development workshops, and works with the Monterey Bay Aquarium both as a volunteer and as a reviewer of science content. Dr. Matsumoto has served on the National Academies of Sciences, Engineering, and Medicine's Ocean Studies Board (2008–2013) and the Board of the National Marine Educators Association (2010–2016), was awarded the QuickScience Ocean Science

Leadership Commitment to Education Award, and is an Association for the Sciences of Limnology and Oceanography Fellow. He has served on a number of review boards for the National Science Foundation, the National Oceanic and Atmospheric Administration, Gulf of Mexico Research Initiative, and the National Academy of Sciences and does his best to spend more time in or on the ocean than on travel.

Pamela S. Soltis (NAS) is a distinguished professor and a curator in the Florida Museum of Natural History and the director of the Biodiversity Institute at the University of Florida (UF). She serves on the executive committee of the UF Genetics Institute and on several committees of the museum and the Department of Biology and has recently served on the UF Graduate Council. She is the director for research at Integrated Digitized Biocollections, the National Science Foundation–funded national center for digitization of biodiversity collections, where she works with the collections community and biodiversity scientists from around the world to develop and promote the use of herbarium specimens (and other natural history collections) in innovative research. She is the president of the American Society of Plant Taxonomists (ASPT) and has served ASPT on the Council (1993–1996), on the Honors and Awards Committee (1993–1995; Chair, 1995), as a Cooley Award Judge (several years; Chair, 1995), and as a reviewer of manuscripts for *Systematic Botany*. She is also the president of the International Society for Phylogenetic Nomenclature and has served her profession as the president of the Botanical Society of America; the president of the Society of Systematic Biologists; a council member for the Society for the Study of Evolution, the International Society for Phylogenetic Nomenclature, and the American Genetic Association; and an associate editor of numerous journals (currently, Board of Reviewing Editors, *Science*; consulting editor, *The Plant Cell*; previously, associate editor for *Systematic Biology*, *Evolution*, *Molecular Biology and Evolution*, *Molecular Phylogenetics and Evolution*, *Taxon*, *Journal of Evolutionary Biology*, *Conservation Biology*). She has received several awards for her contributions to the study of plant diversity, most notably the International Prize in Botany (Physiographic Society of Lund, Sweden), the Asa Gray Award (American Society of Plant Taxonomists), the Darwin-Wallace Award (Linnean Society of London), and the Botanical Society of America's Merit Award, all jointly with Douglas E. Soltis. Dr. Soltis received a B.A. in biology from Central College (Pella, Iowa) (1980), a Ph.D. in botany from The University of Kansas (1986), and an Honorary Doctorate of Humane Letters from Central College (2017). She is a member of the National Academy of Sciences and the American Academy of Arts & Sciences.

Barbara M. Thiers is currently a vice president and the director of the William & Lynda Steere Herbarium of the New York Botanical Garden, where she has been since 1981. From 2014 to 2017, Dr. Thiers oversaw the Garden's research division and continues to serve in an advisory role to the chief executive officer and chief operations officer of the institution today. She earned her Ph.D. in botany from the University of Massachusetts. Her research area is the systematics of the Lejeuneaceae, a family of leafy hepatics. Since becoming the director of the herbarium, Dr. Thiers has managed and raised funds for the facility, which contains approximately 8 million specimens. The Steere herbarium is among the three largest herbaria in the world and the largest in the Western Hemisphere. Since 2008 she has managed the online resource Index Herbariorum, which is a directory of the approximately 3,000 herbaria worldwide. In 2010, Dr. Thiers served on the National Science Foundation (NSF)-funded committee to develop the Networked Integrated Biocollections Alliance strategic plan for the digitization of natural history collections in the United States. This plan led to the establishment of NSF's Advancing Digitization of Biodiversity Collections funding program (2011 to present). Currently, she serves as a member of the External Advisory Committee for Integrated Digitized Biocollections, and the Biodiversity Collections Network Advisory Committee. She is also currently the president of the Society for the Preservation of Natural History Collections, the vice president of the Natural Science Collections Alliance, and a member of the external advisory committee for the Harvard University Herbaria & Libraries.

STAFF

Audrey Thévenon is a program officer for the Board on Life Sciences at the National Academies of Sciences, Engineering, and Medicine, where she also serves as the managing editor of the *Institute for Laboratory Animal Research Journal*. Since joining the National Academies, Dr. Thévenon has supported collaborative regional and international activities at the intersect of infectious disease research and policy decision specifically aimed at promoting transdisciplinary research in global health. Dr. Thévenon has been involved in activities that support the Department of Defense's programs to counter biological threats and to inform about the potential risks and benefits of gain-of-function research, and supported a study on gene drive research in non-human organisms. Currently, she leads a One Health fellowship program in Pakistan and the Response and Resilient Recovery Strategic Science Initiative launched to run prospective crisis management scenarios related to coronavirus disease 2019. Prior to joining the National Academies, Dr. Thévenon completed a postdoctoral fellowship at the University of Hawaii in placental pharmacology, followed by another fellowship at the Uniformed Services University of the Health Sciences in Bethesda, Maryland, working on two President's Emergency Plan for AIDS Relief–funded HIV-Malaria projects in collaboration with Nigeria and Kenya. Dr. Thévenon has a Ph.D. and an M.S., both in biology from Georgetown University with a specialization in tropical medicine and immunology, as well as an M.S. in cell biology and physiology from the Université de Rennes 1 in France.

Keegan Sawyer is a senior program officer for the Board on Life Sciences at the National Academies of Sciences, Engineering, and Medicine. Her work addresses a wide range of research, policy, and communication questions across the broad spectrum of life science disciplines. She has a special interest in the interplay of environmental conditions and human health, ecosystem health, and public engagement in science. Dr. Sawyer is the director of the National Academies' Standing Committee on the Use of Emerging Science for Environmental Health Decisions. She recently served as the project director for the Committee on Gene Drive Research in Non-Human Organisms: Recommendations for Responsible Conduct and the Committee on Value and Sustainability of Biological Field Stations, Marine Laboratories, and Nature Reserves in 21st Century Science, Education, and Public Outreach. She is committed to fostering discussions about research infrastructure, collaborative environments, and public engagement in science to support healthier people and planet. Dr. Sawyer holds a B.S. (1999) in environmental biology from University of California, Davis, and an M.S. (2002) and a Ph.D. (2008) in environmental sciences and engineering from the University of North Carolina Gillings School of Global Public Health.

Jessica De Mouy is a senior program assistant for the Board on Life Sciences at the National Academies of Sciences, Engineering, and Medicine. She worked on the *Report of the Committee on Proposal Evaluation for Allocation of Supercomputing Time for the Study of Molecular Dynamics, Tenth Round* (2019). Additional projects include workshops for the Standing Committee on the Use of Emerging Science for Environmental Health Decisions and the Committee on Assistance to the U.S. Fish and Wildlife Service on Taxonomic Studies of the Red Wolf: A Review of Applications to Carry out Research and Development of a Research Strategy. She holds a B.A. (2018) in sociology from the University of Maryland, College Park.