Biodiversity Collections, Data, and COVID

BETH BAKER

From virus detection to innovative teaching, collections are key to research, education.

At this writing, more than 24 million cases of coronavirus disease (COVID-19) have been tracked around the globe, and there have been more than 825,000 deaths. The numbers are increasing dramatically, especially in the United States and the global South. Yet researchers still do not know where SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) originated or how it spread. An unidentified species of Chinese horseshoe bat is the current best guess, with an intermediary mammal host allowing the virus to jump to humans, where it evolved to become a novel virus.

A small carnivore, the masked palm civet, was implicated in the first SARS-CoV outbreak in 2003 in China. The civets also may have become infected by horseshoe bats. But when SARS-CoV-2 emerged 17 years later, virologists were still in the dark.

While scientists race to develop a vaccine, far less attention is being paid to determining the virus's provenance. Only with that knowledge, say experts, can policy changes be developed to address capturing, selling, preparing, and consuming wild animals that are the likely source of recent pandemics. And the key to identifying a virus and understanding its evolution may well lie hidden in biodiversity collections.

"The way you get ahead of those kinds of disease epidemic initiations is to understand everything about how the crossing over actually happens and where it comes from," say John Bates, curator and section head of life sciences at The Field Museum in Chicago. "We see our collections and associated databases as more valuable than ever."

Dedan Ngatia (then at Karatina University and now at the University of Wyoming) and Adam Ferguson collect blood samples from anesthetized zorillas (Ictonyx striatus) in Laikipia County, Kenya, in 2015. Photograph: Jennie Exley.

His colleague at the museum, Adam Ferguson, had already been working with Canadian researchers to try to isolate DNA viruses from historic museum specimens. When the COVID pandemic hit, they shifted their focus to bat specimens, collected from around the world since 1896. "That provides an amazing vault of information," says Ferguson. "We didn't have to mobilize funds or efforts for fieldwork to go collect new bat specimens." Ferguson hopes to learn more about the evolutionary history of these viruses, as a critical part of the story of how they jump to humans. Only with such knowledge, he says, can there be hope of averting global pandemics.

BioScience XX: 1–7.
doi:10.1093/biosci/biaa093

Pandemic shows vulnerability of nation's collections

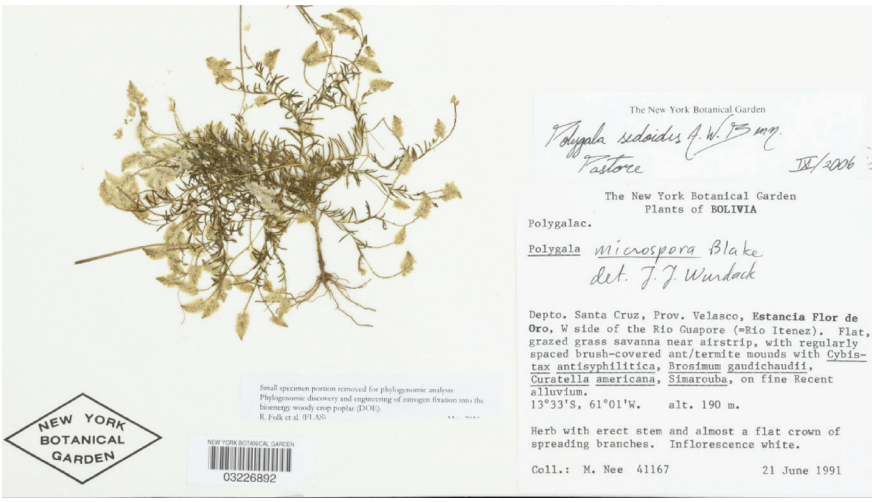
Disease detection is one of many vital missions that biodiversity collections play, along with environmental assessment, online education, and detection of crop disease and potential bioterror organisms, among many other uses. But the COVID-19 pandemic has shown that collections are vulnerable, as museum and university budgets are cut, research is paralyzed, and the specimens themselves are at risk when staff are not there to maintain them.

In April, a survey of 390 collections professionals, conducted by the American Institute of Biological Sciences, Biodiversity Collections Network, Natural Science Collections Alliance (NSC Alliance), and Society for the Preservation of Natural History Collections, uncovered respondents' concerns about the effects of pandemic closures. Among the findings: Nearly two-thirds were worried they would not be able to fulfill research requests. Nearly half were concerned about funding cuts and potential staff losses. And although most had designated essential workers to monitor their collections, less than 30% had staff who were able to check for pests, which are a major threat to specimens.

"Our big concern is pest damage," says Barbara Thiers, director of the herbarium at the New York Botanical Garden, which closed in mid-March but had a limited opening in August. "I didn't get worried until the weather warmed up, which is much more conducive to critters." Added to that is hurricane season and the threat of storm damage.

Collections such as hers that are associated with gardens or museums have been hard hit financially as the public was kept out and entry fees disappeared overnight. A major blockbuster spring show at the botanical garden was postponed, and she was worried that other public programs may be canceled in coming months.

As NSC Alliance president, Bates hears from curators and collections managers around the country. "For freestanding museums, we were able to get some support from government, but we are definitely concerned about the next 6 months," he says. Research by graduate students and postdocs, as well as grant money, all are tied to deadlines that might not be met.



This specimen from the New York Botanical Garden's herbarium was used by researchers at University of Florida to study nitrogen fixation, with a goal of introducing nitrogen-fixing genes into bioenergy crops. Photograph: C. V. Starr Virtual Herbarium (http://sweetgum.nybg.org/science/vh).

Pam Soltis, who coheads the Laboratory of Molecular Systematics and Evolutionary Genetics at University of Florida's Museum of Natural History, says their institution depends on state funding. "Florida has a balanced budget law so that if less money comes into the state, then less goes out to state-supported agencies and institutions," she says. She worries about potential staff cuts.

Meanwhile, researchers around the world have been unable to access voucher specimens. "There's a huge impact on research activity," says Scott Miller, chief scientist of the Smithsonian Institution. "Staff and visitors have not had access to the physical collection," other than the core group who maintains the building's environment and security and checks the collections. With field research at a standstill, "2020 is already becoming known as the year of missing data" among ecologists, Miller says.

Although the Smithsonian's collections are well supported, many small collections scattered around the country may be at risk. A June survey of US museums showed that 16% feared they might permanently close without additional financial relief; of the respondents, 2% were arboretums or botanical gardens and 4% were natural history or anthropology museums. "There are collections that had to close during this time, and they were not able to continue environmental monitoring," says Miller. Collections can be damaged by air conditioning failure, for example. "Dry collections like herbaria are really vulnerable to high humidity, which degrades their DNA and promotes the growth of fungus," he says.

Until museums and universities fully reopen, the fate of collections is unknown.

The ever-expanding role of collections

Collections in the United States are home to some 3 billion specimens. The concern over their safety and viability has grown over the last decade even as their scientific and economic contributions expand. New technologies, from genomics to isotope analysis, allow scientists to use historical specimens to uncover a vast trove of information about life on Earth.

NSF research fellow Nicolas Dowdy with the Milwaukee Public Museum uses specimens of tiger moths (Erebidae: Arctiinae), to study the evolution of sound production among these moths. Images: Nicolas Dowdy.

Among them: Specimen data is used to identify potentially devastating invasive species at ports of entry. The data are also a valuable tool in detecting and securing dangerous biological agents, as was discussed in a 2019 workshop held by the National Defense University. Even aviation safety has been aided by collections, as was demonstrated in 2009, when researchers used DNA to determine that migrating Canada geese caused the crash of a US Airways jet. "Integrating this information with bird migration patterns, bird-detecting radar, and bird dispersal programs at airports can minimize the risk of such collisions in the future," according to a study in *Frontiers in Ecology and the Environment*.

Brown University researchers used herbarium specimens to measure changes in heavy metals, comparing nineteenth century specimens with those collected in 2015. They found that copper and zinc levels were unchanged, while lead had decreased over time. "These specimens can be a valuable part of environmental science research," they wrote in a 2017 issue of *Applications in Plant Sciences*.

In agriculture, specimens yield insights into crop diseases such as the blight that caused the Irish potato famine or today's rust fungi diseases. "We have millions and millions of specimens digitized," says Thiers. "You can correlate the diseases with where and how things have changed over time. With climate change there are concerns that pathogens that aren't serious now could become more serious with changes in rainfall, climate, soil moisture. We can look at the specimens and use genomic techniques to understand what genes are activated in plants and in parasites."

Investing in biodiversity collections offers a "big payback," says Joseph Cook, curator of mammals at the Museum of Southwestern Biology at University of New Mexico. In addition to demonstrating to the public the "incredible diversity of what's on this planet," he says, collections help to answer big societal questions, related to climate change, agriculture and food security, and invasive species. "These are all issues of change through time. You have to have baselines in the past, and the best place to see that is in museums." For example, University of Florida researchers used specimens from Thiers's collection to try to find the genes responsible for nitrogen fixation.

Even biologists may not fully appreciate the value of collections, says Cook. "The taxonomy and systematics traditionally held in museums are fundamental to a lot of other areas of biology," he says, and collections need to do a better job of telling their story. "Museums have evolved a lot over the last 30 years. We serve more biodiversity data than anyone."

Global access to specimens is made possible by the robust effort to digitize the nation's biodiversity collections, especially critical during this time when field research came to a halt. Beginning in 2011, the National Science Foundation (NSF) launched a 10-year, $100 million push to support a digitization infrastructure.

Since then, the Integrated Digitized Biocollections (iDigBio) network has amassed over 120 million records from the nation's collections. Some 100,000 of these specimens are also linked to GenBank, the genetic-sequence databank at the National Institutes of Health (NIH). MorphoSource, housed at Duke University, allows researchers, collections curators, and the public to find and share 3D images of biological specimens.

"We have huge numbers of images now, and they can be used for all sorts of other data, things we couldn't even imagine in 2011," says Soltis who leads iDigBio's research coordination and scientific community outreach. "Exciting discoveries in computer science and machine learning technologies are enabling us to get information relatively easily, and this is just the beginning."

There has been rapid progress in the ability to link together different types of data, says Soltis. Data from

a single voucher specimen may be used by many researchers for different purposes. "The fact that we can have DNA-sequenced data in GenBank that are linked to a voucher specimen in a museum somewhere," she explains, "and linked to an image somewhere else that has traits extracted from that image, and with links to stable isotopes that have been used for some other study, and all those pieces of information can be linked through unique identifiers—that's an amazing view of biodiversity we've never been able to have before."

At the New York Botanical Garden's herbarium, roughly half of the 8 million specimens have been digitized. "We considered it a major feat to do 10,000 specimens a year," says Thiers. "Now we do 400,000 to 500,000 specimens a year. We used to have someone pick up a specimen, enter the data, maybe take a picture. Now we image thousands of specimens and our staff and volunteers can look at the images and enter the data. We've built in tons of efficiencies that are now widespread in the [collections] community."

During the down time of the pandemic, collections staff worked from home to improve the quality of their digitized specimens. Many specimens that have been imaged have not yet been georeferenced or had the data transcribed. Other specimens have mistakes in the data. At the Milwaukee Public Museum, for example, employees set up work stations at home to continue the task of scanning slides of insects. The images can then be shared with students and volunteers to transcribe the data.

"Not only are we continuing to digitize from home, we're also able to take a step back from the data transcription and imaging and georeferencing and examine how our collections data are being used," says Jennifer Zaspel, research curator and head of zoology. Her team is conducting a special assessment of the data and looking for patterns and species interactions that can be categorized. "It's been exciting to go into the records that we've made available to see that so many are being downloaded and used."

When the Texas A&M University Insect Collection was preparing to shut down due to the pandemic, student employee River Martinez, shown here, and assistant curator Karen Wright went to work imaging as many specimens as possible. Students were trained to work on the specimen image data from home. When the shutdown dragged on, Martinez went to the campus and retrieved additional specimen slides to take home, to work on more images. Photograph: River Martinez.

Nicolas Dowdy, a postdoctoral NSF research fellow with the Milwaukee Public Museum, works with tiger moth specimens to study the morphology of the moths' sound-producing organs. He extracts defensive chemicals, which remain after death, from the specimens to learn about the chemicals' diversity and perhaps their host plant. He also extracts DNA to reconstruct the evolutionary history of the lineages and their traits and uses the label data to understand the traits' spatial distribution. "The size and scope of the project is such that it couldn't feasibly be done without utilizing museum material," wrote Dowdy in an email. "Well, at least not without a lifetime or two's worth of work."

Using collections in public health

With some 75% of emerging infectious diseases coming from animals, collections can play a critical role in tracking outbreaks before they become pandemics. That was demonstrated in 1993 with hantavirus in the Four Corners region of the US Southwest. When Navajo Nation people began dying from a rapid and severe pneumonia, a team of public health experts fanned out to find the cause. They enlisted the Museum of Southwest Biology at the University of New Mexico, which had preserved specimens of rodents and frozen tissue samples. "They're available as a reference library for virologists to screen for different kinds of pathogens," says Cook. "They were able to screen our large historic series of deer mice and associated samples which showed that this pathogen was not new in 1993." It also was discovered that higher rainfall from El Niño events leads to an explosion of deer mice populations, which then enter homes on the Navajo Nation and throughout the Southwest and spread hantavirus.

Since then, Cook has developed ongoing relationships with other virologists. "Over the last 15 years I've been working with a hantavirologist from Hawaii and others, and

Schuyler Liphardt, doctoral student at the University of New Mexico, wearing personal protective equipment as he conducts field work. Liphardt studies the evolution of hantaviruses carried by rodents and shrews in North America. The deer mouse (right) was the carrier of hantavirus in the 1993 epidemic on the Navajo Nation and elsewhere in the Southwest. Photographs: Joseph Cook.

our paradigm has been completely revised," he says. "It was originally thought that hantavirus was a rodent-borne disease, but using natural history collections of shrews, moles, and bats, we found they all have their own [species of] hantavirus." This work has helped inform researchers about major hantavirus outbreaks throughout the Americas.

Ferguson, of the Field Museum, works with Maasai Mara University, Kenya's Zoonotic Disease Unit, the International Livestock Research Institute, and other organizations at the interface of parasites, humans, livestock, and wildlife. In particular, he studies genets, small, common carnivores that carry diseases, such as rabies. They, together with other small carnivores such as mongooses, are also sold at markets and sometimes eaten by people. "I work really closely with pathobiologists, especially veterinarians and epidemiologists," Ferguson explains. "I'm not a virologist, but I'm an expert in small mammal ecology and behavior and evolution."

Working across disciplines, Ferguson collaborates with a multitude of scientists to screen specimens or blood samples for particular pathogens of interest. Similarly, a virologist might contact Ferguson in order to find a tissue sample. Still, the long-time role of collections experts in pathogen surveillance often is overlooked when a pandemic strikes. "Natural history museums have been left out of the conversation, especially out of One Health," he says. The Centers for Disease Control and Prevention's One Health collaborative works at all levels, from local to international, to examine the relationships among humans, plants, animals, and their shared environments.

To strengthen relationships among the virology, disease ecology, and collections communities in the wake of the COVID-19 pandemic, museum faculty and staff initiated the ViralMuse task force, supported by iDigBio and others. For example, mammal cultures and virus cultures may be housed in different collections. "From a given mammal, you might have skin tissue and you might well have pathogens in another collection," says Thiers. "If we could search all of these databases at once, we could be much more facile in

Before the pandemic, Ashley Linderman, Abdulrahman Alhilal, and their classmates in the Form and Function course at Central Michigan University worked in a lab. The lab has since been adapted as an online lab experience. The materials, many of them using collections images and data, are available at http://biodiversityliteracy.com. Photographs: Adam M. Sparkes with University Communications at Central Michigan University.

how to respond to questions of where does a virus come from."

Another project that could predict future disease outbreaks is under way at the Milwaukee Public Museum's zoology department. "Our current [digitization] project is heavily focused on collecting baseline distribution data for arthropod organisms that transmit pathogens, or have the potential to transmit them," says Zaspel. "The idea is to transcribe label data from those species that are known to harm humans and livestock and wildlife. Not only are we transcribing the locality and collection information from the specimens, but the specimens often come with information about the host, so we're transcribing that data. In the broadest sense, what we hope to have at the end of the project is a database that brings all of this information about where and when these species occur in the United States, together in one place."

This would allow researchers to predict where the next pandemic might strike. "If we put in the investments now to build the collections and do the vouchering and databasing, we'll be in a much better position in the future," says Bates.

Digitized collections save the day for educators

With the sudden lockdown from the pandemic, university faculty had to move their courses online and scramble to continue to provide meaningful hands-on experiences for their biology students.

At Central Michigan University, professor and herbarium director Anna Monfils was already leading an effort to incorporate natural history collections into undergraduate science classes. As principal investigator for Biodiversity Literacy in Undergraduate Education–BLUE Data Initiative project, a Research Coordination Network in Undergraduate Biology Education funded through the NSF, Monfils works with collaborators to develop innovative ways for students to access and use biodiversity collections.

When COVID-19 closed the campus, Monfils and her team went to work adapting the collections-based educational materials for online delivery. Realizing educators may be overwhelmed by a flood of resources and limited in their time, BLUE ran an implementation blitz to introduce a curated set of collections-based educational materials and help teachers use the materials in an online format.

Monfils is now working on a lesson on bees that asks students to go outside and take a cellphone image of a bee using an online application and citizen science platform, iNaturalist. They post pictures to iNaturalist that can in turn be viewed and verified by citizen scientists. "They're interacting with the scientific community," says Monfils. "Not only does this student have this experience, they're thinking globally and are contributing to a public data set and [informing] the known distribution of an insect. Science is really being done. Later when they encounter a data set, they have an understanding of where data come from." Plus, she adds, it is more fun for students than yet another class sitting in front of a computer monitor.

"We are currently producing and modifying materials for educators to move to online undergraduate research experiences with BLUE materials and are developing a potential online undergraduate symposium or conference to allow students to present these materials virtually," says Monfils.

At Widener University in Chester, Pennsylvania, associate professor of biology Janice Krumm and colleagues received a $200,000 RAPID grant from the NSF to develop online research experiences for more than a thousand undergraduate students around the nation. Krumm, along with Elizabeth Shea, curator of mollusks at the Delaware Museum of Natural History, had already incorporated digitized collections into coursework, along with in-person museum research. "When we moved online in the spring, student projects shifted to focus more heavily

on the digital data and the students were able to proceed with really great research projects," says Krumm.

During the spring class, students focused on Xenophoridae (carrier shells), a family of sea snails. "They attach objects to their shells, and the students wanted to explore why they do this," Krumm explains. Originally the students were taking measurements using Shea's mollusk collection. But when the pandemic hit, they focused on three species found around Florida and the Caribbean. Using digitized data, they studied how the snails' distribution matched up with marine ecoregions and water depth. Without access to collections data, "It would have turned into a research paper based on the literature," says Shea. "It would have been dialed way back."

Students are also able to add georeferencing to specimen data and learn about systematics. "As a science educator, I see collections as a tremendous opportunity to do active engagement with research," says Carly Jordan, associate professor at George Washington University who joined Krumm and Shea in developing the online materials. "Educators are going to change how we think about our courses from here on out. Trying to identify these opportunities that are meaningful, but not in-person, is extremely important."

She adds that digitized collections offer smaller universities and community colleges that do not have research labs a way to access resources they otherwise would not have. "It brings up important and timely issues of equity in STEM, and improves access to research experiences for underrepresented groups," she says.

The coursework has the added benefit of expanding the universe of support for biodiversity collections. "iDigBio has been great over the years, always looking for ways to reach new audiences of biologists, including conservation, and ecology researchers," says Shea. The team also is enthusiastic about the chance to share their work with colleagues across disciplines and with institutions around the country.

The Division of Genomic Resources, Museum of Southwestern Biology at the University of New Mexico, now preserves nearly 600,000 cryovials of tissue from more than 290,000 voucher specimens. Photograph: Joseph Cook.

Moving forward

In the coming years, collections professionals anticipate that digitization will continue to rapidly expand access to specimens and that new technologies will allow specimens to be used in novel ways. To encourage these developments, the collections community is coming up with standardization criteria, allowing data to be more easily linked.

Using collections, researchers would be able to more accurately predict future human pandemics, as well as crop diseases or pest infestations. The historic environmental conditions captured in specimens will continue to shed light on how organisms and ecosystems will respond to a changing climate. "We can begin to predict where the next problems might come from," says Thiers.

As was described in the January issue of *BioScience* (dx.doi.org/10.1093/biosci/biz140), she and other collections experts are promoting the Extended Specimen Network initiative to significantly expand the value of specimens.

"The Extended Specimen Network overcomes the silos and makes it easier for people to find and use our collections," says Bates. He compares it to NIH's GenBank, but for specimen-based phenotypic data. "[The Extended Specimen Network] is equivalent to inventing the computer or the cell phone; it's that basic," he says. The goal is to have a stable archive of data, freely accessible to researchers around the world, with a reliable source of funding. "This COVID example shows clearly that humans are not separated from biodiversity and likely never will be. We can learn so much."

Beth Baker (bbaker@aibs.org) is a freelance journalist and the Features editor of BioScience.