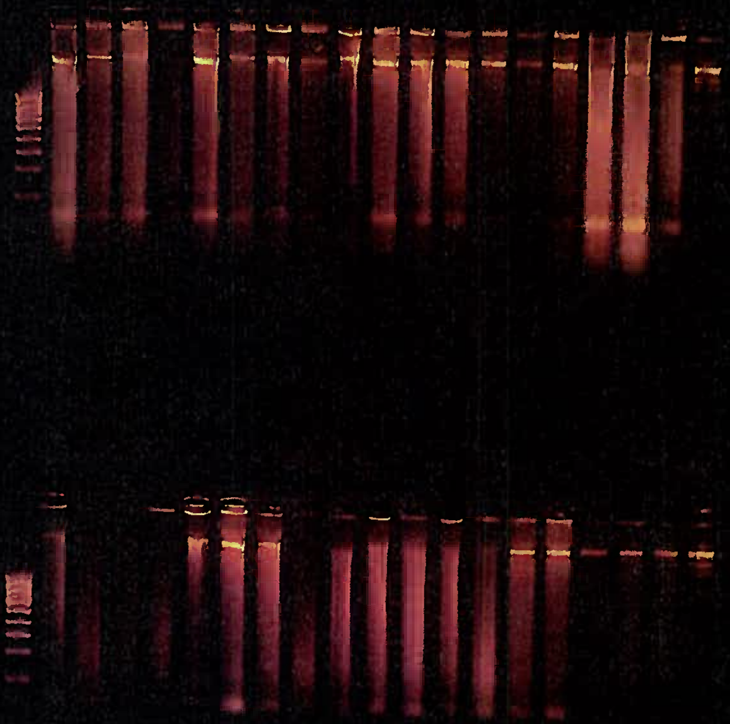
DNA Banking for the 21st Century



Wendy L. Applequist and Lisa M. Campbell, Editors

A New iDigBio Web Feature Links DNA Banks and Genetic Resources Repositories in the United States[1]

Grant T. Godden[2] and Pamela S. Soltis[3]

ABSTRACT

Collections of genetic resources have served the needs of molecular biologists for many decades, providing valuable sources of materials for molecular studies of biodiversity and facilitating long-term archival of samples from published studies. However, the needs of the research community are rapidly changing and contributing to an increasing demand for high-quality genetic resources. Curators and researchers in the United States (U.S.) have recently engaged in an active dialogue about the future of these collections, as well as about best practices and standards for collecting and preserving new materials. The discussion has reemphasized a previously identified need for networked data on genetic resources to better link samples housed in U.S. collections with research consumers. However, there are no comprehensive lists of DNA banks, genetic resources repositories, or tissue collections available, and significant gaps remain in our knowledge about the types of genetic resources that are available as well as their physical localities. In this paper, we describe a newly available web feature developed and maintained by Integrated Digitized Biocollections (iDigBio) that is helping to compile the first comprehensive list of genetic resources collections in the U.S. Our new tool provides information on a diverse set of collections and enables rapid searching of multiple collections.

KEY WORDS: DNA bank, genetic resources, iDigBio, natural history, tissue collections

Recent advances in molecular research techniques and sequencing technology, as well as their widespread adoption in the biological research community, are increasing the demand for suitable tissues for genomic, transcriptomic, and other contemporary molecular research applications. This rising demand for samples has prompted community-wide discussions—engaging both museum curators and researchers from many disciplines in the biological sciences—regarding current and anticipated research needs, and the utility of available resources currently housed in biorepositories in the United States (U.S.) and abroad (e.g. U.S. Workshop on DNA Banking, Missouri Botanical Garden, January 2013; Capturing Genomes Workshop, Global Genome Initiative, National Museum of Natural History, Smithsonian Institution, April 2013).

Specimens housed in natural history collections have served as important sources of tissues for molecular studies of biodiversity for many decades. However, these specimens were carefully preserved for morphological study, and destructive sampling for molecular research potentially decreases their value for other important areas of research. Moreover, the nucleic acids derived from specimen tissues are sometimes suboptimal for contemporary molecular research; commonly employed methods for specimen preparation were not designed for long-term preservation of nucleic acids, and the quality of genomic DNA derived from these tissues may produce undesirable results.

The establishment of DNA banks or genet-

---

[1] Integrated Digitized Biocollections (iDigBio) is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

[2] Department of Biology, University of Florida, Gainesville, Florida, U.S.A. g0ddengr@ufl.edu.

[3] Florida Museum of Natural History, University of Florida, Gainesville, Florida, U.S.A. psoltis@flmnh.ufl.edu.

ic resources repositories has not only helped mitigate potential damage to museum specimens, but has also helped service the demand for tissues suitable for molecular research. These important biological collections comprise samples that were prepared with the primary goal of preserving either nucleic acids extracted from field-collected samples (prior to museum voucher preparation) or unprocessed field-collected tissues in long-term storage conditions. Consequently, these samples are usually better suited for molecular research than tissues derived from museum specimens themselves. Nevertheless, the methods by which some tissues are preserved at the time of collection, or extracted post-collection, may still render samples unsuitable or suboptimal for some types of molecular work, especially given the requirements of next-generation sequencing and assembly technologies. For example, some genome-scale research applications currently require large yields of starting DNA of high molecular weight, and transcriptomic work necessitates flash freezing of tissues in liquid nitrogen at the time of collection, followed by immediate and permanent storage at ultracold or cryopreservation temperatures to prevent degradation of ribonucleic acids (RNA). Therefore, it is no surprise that samples maintained as part of DNA banks or genetic resources repositories may vary greatly in their quantity, quality, and corresponding utility for contemporary research (see Neubig et al., this volume). Depending on the molecular approach, a biological researcher may require more detailed information about the method of field collection, preservation, transport, and storage conditions when making decisions about the suitability of a particular sample for his or her study.

Very recently, the biodiversity research community has developed an interest in designing new, organism-specific best practices and standards for sample collection, transport, and storage (e.g. Wong et al., 2012; Amato, 2013; Zimkus, 2013) in order to: (1) ensure optimal preservation of nucleic acids and other molecules in tissues of diverse biological samples and (2) maximize the long-term utility of curated genetic resources for a broad range of molecular research applications. Moreover, concerted efforts are already underway to develop Darwin Core Archive DNA and tissue data standards (e.g. Dröge et al., 2013; Flemons et al., 2013) that facilitate both the capture of sample metadata (including detailed histories and quality information) and the development of searchable databases. The resulting products of these efforts will better serve the current and future needs of the biodiversity research community, allowing individual researchers to source suitable samples for a broad range of molecular research applications. However, the absence of a comprehensive list of DNA banks and genetic resources repositories continues to represent a major roadblock to this endeavor.

ESTABLISHING A NETWORK OF DNA BANKS AND GENETIC RESOURCES REPOSITORIES

The DNA banking movement began in the late 1980s, and collective efforts were made throughout the 1990s to establish an international network of repositories whose collective mission included the long-term preservation of biological tissues and distribution of nucleic acid extracts for molecular studies (Adams, 1988; Adams et al., 1992). By 1992, 31 institutions participated in the newly established network "DNA Bank-Net" (Adams, 1997). Today, this network is known as the DNA Bank Network and is maintained as a node of the Global Genome Biodiversity Network (GGBN)(http://www.dnabank-network.org/).

A comprehensive list of DNA banks and genetic resources repositories does not currently exist. In fact, with the exception of the limited resources available in the DNA Bank Network, there are currently few online resources available to aid researchers in sourcing suitable genetic resources for current or future scientific studies or identifying potential repositories for archival of tissues and extracts associated with active and previously published research. The lack of information about genetic resources poses many important questions, including: (1) how many collections exist and where are they located; (2) how are the genetic resources collected, prepared, stored, and curated within institutions; (3) how many and what types of samples are included in each collection; and (4) how are available resources made known to the scientific community?

It is common knowledge within the systematics community that many natural history collections around the world have already established or are beginning to establish curated collections of nucleic acid extracts (DNA or RNA) and preserved tissues for molecular studies. These collections constitute DNA banks or genetic resources repositories, but may not be labeled as such. For example, frozen tissue collections and DNA from vouchered specimens are sometimes maintained as a part of departmentalized or independent institutional collections (e.g., mammal or bird collections), while others are curated within centralized repositories that include samples from a diverse range of organisms (Zimkus, 2013; see also Zimkus and Ford, this volume).

The terms "DNA bank" and "genetic resources repository" currently lack a specific definition, and there is no apparent pattern regarding the preferential use of either term in representing a curated collection of genetic resources (if a term is used to describe them). The terms currently encompass a wide variety of different collection types, and this poses considerable challenges for researchers who wish to identify existing collections and repositories with Internet searches. The generalized terms, "DNA bank" or "genetic resources repository," are currently used to describe *ex situ* living and germplasm collections (Rice et al., 2006), seed banks and gene libraries (Given, 1994; Migliani, 1998), extracted nucleic acid collections (Mattick et al., 1992; Adams, 1997; Rice et al., 2006), silica-dried tissue collections (Adams et al., 1999), frozen tissue collections (Adams, 1993), and cell lines and culture collections. As defined here, a DNA bank is a curated collection of nucleic acid extracts, whereas a genetic resources repository is any perpetually maintained collection of curated tissues, nucleic acid extracts, cell lines, or living cultures; both are preserved in a long-term storage facility and available to the scientific community for molecular research applications.

As the National Resource for Advancing Digitization of Biodiversity Collections (ADBC), Integrated Digitized Biocollections (iDigBio) is well positioned to solicit, compile, and serve information about genetic resources via its website (http://idigbio.org). As of early 2013, we began researching and compiling information about DNA banks and genetic resources repositories in the U.S. Based on our preliminary data, we were able to develop a simple organizational scheme to display descriptive information about individual collections of genetic resources and to link commonly shared attributes among collections that may be of importance to the research community. We kept in mind many of the sample suitability issues described above and tried to capture important details regarding preservation and storage using linked keywords. Lastly, we developed a user-friendly web feature and, in March 2013, solicited the help of the U.S. Workshop on DNA Banking participants to populate our new web resource with information about known collections. We also advertised the launch of our new resource in May 2013 as an iDigBio news feature and solicited community-wide input via our extensive network of participating Thematic Collections Networks and iDigBio subscribers.

As of January 2014, 52 U.S. collections of genetic resources are included in iDigBio, representing a variety of collection types and a broad sample of biological diversity. We recognize that many more collections likely exist throughout the U.S., and we welcome input from the broader scientific community to aid our continued efforts to identify and supply information for all U.S. collections of genetic resources.

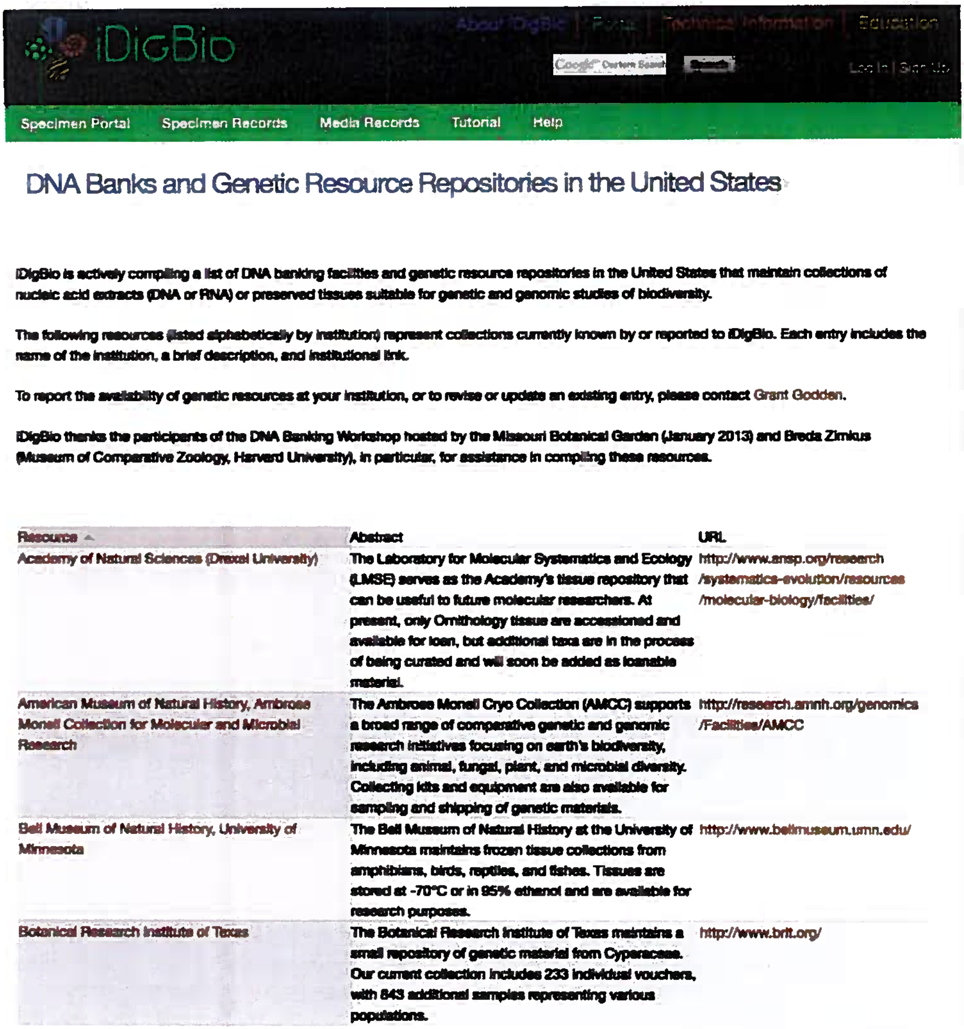
ABOUT THE IDIGBIO RESOURCE

The primary interface for our newly developed iDigBio web feature provides an alphabetical listing of all DNA banks and genetic resources repositories in the U.S. known to iDigBio (https://www.idigbio.org/genetic-resources). Brief details about each collection are presented in tabular format, including the resource name or title of the collection, an abstract briefly describing the collection and its holdings, and an active URL to redirect visitors to each resource (Figure 1). By clicking on a listed resource name [e.g., "Florida Museum of Natural History, Genetic Resources Repository (University of Florida)"; see Endara et al., this volume], visitors to our portal are redirected to the corresponding record for the DNA bank or genetic resources repository, which is maintained as a bibliographic entry in our database. Each database record contains expanded information and indicates the year of publication (or most recent update) as well as a list of keyword links describing the attributes of the selected collection (Figure 2). In our Florida Museum of Natural History example, the keyword links describe the storage conditions (e.g. "-180C" and "liquid nitrogen"), the type of collection (e.g. "centralized repository", "cryogenic collection", and "DNA bank"), the type of samples housed in the collection (e.g. "frozen tissue collection" and "nucleic acid extracts"), the biological organisms represented in the collection as defined by discipline or organismal category (e.g. "botany", "entomology", "herpetology", "ichthyology", "invertebrates", "mammalogy", "ornithology", and "vertebrates"), and the availability of a searchable database (i.e. "online database"); all records are linked in the iDigBio web feature with a common keyword (i.e. "genetic resources"). By clicking on a linked keyword (e.g. "cryogenic collection"), the visitor is redirected to an iDigBio Bibliography page that contains a full alphabetical listing of all records classified with cryogenic storage facilities (Figure 3). By taking advantage of our keyword-based organizational scheme, browsing visitors can easily identify all other collections that share similar attributes of interest and subsequently peruse individual resources via the direct institutional links.

DISCUSSION AND FINAL REMARKS

Resources such as DNA banks and genetic resources repositories represent valuable collections that can aid studies of biodiversity and advance other areas of biological research. However, the general insularity of these collections precludes potentially valuable interactions among the curators who supervise them and limits the use of their samples by the research community.

Without a comprehensive list of repositories and curated collections of genetic resources, it remains difficult to move forward with the establishment of either a U.S.-based or international network of genetic resources. The ultimate goal of such a network is improved access to samples, and any future network should allow researchers to search all collec-



DNA Banks and Genetic Resource Repositories in the United States

iDigBio is actively compiling a list of DNA banking facilities and genetic resource repositories in the United States that maintain collections of nucleic acid extracts (DNA or RNA) or preserved tissues suitable for genetic and genomic studies of biodiversity.

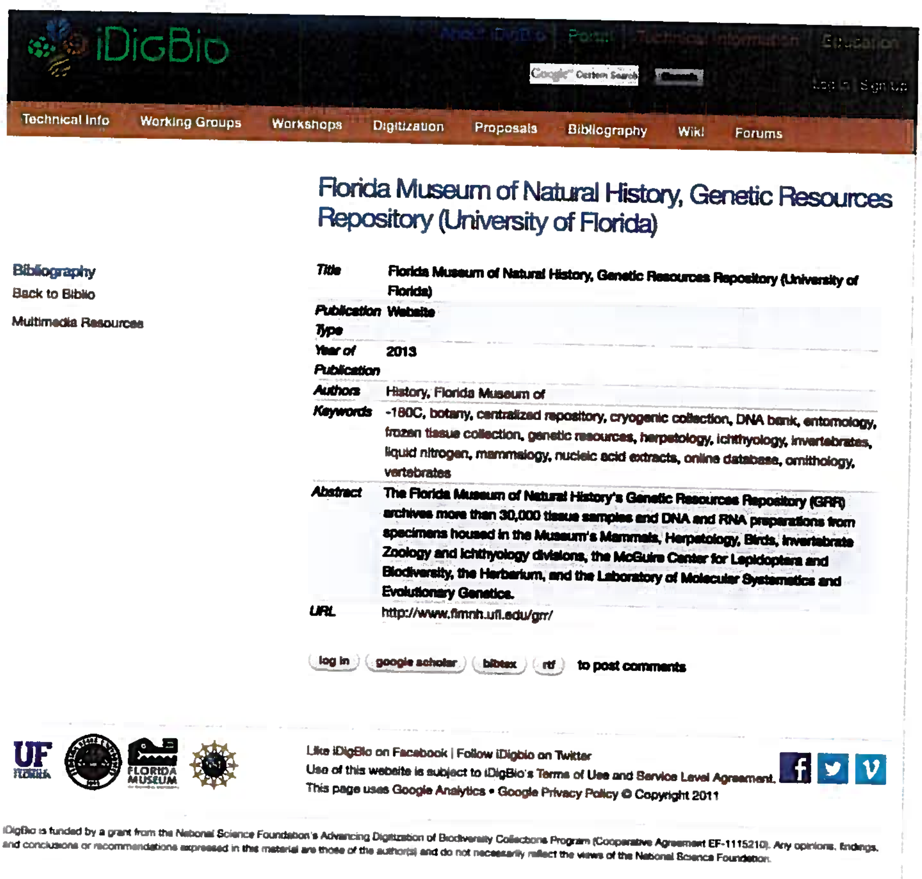
The following resources (listed alphabetically by institution) represent collections currently known by or reported to iDigBio. Each entry includes the name of the institution, a brief description, and institutional link.

To report the availability of genetic resources at your institution, or to revise or update an existing entry, please contact Grant Godden.

iDigBio thanks the participants of the DNA Banking Workshop hosted by the Missouri Botanical Garden (January 2013) and Breda Zimkus (Museum of Comparative Zoology, Harvard University), in particular, for assistance in compiling these resources.

Resource ▲	Abstract	URL
Academy of Natural Sciences (Drexel University)	The Laboratory for Molecular Systematics and Ecology (LMSE) serves as the Academy's tissue repository that can be useful to future molecular researchers. At present, only Ornithology tissue are accessioned and available for loan, but additional taxa are in the process of being curated and will soon be added as loanable material.	http://www.ansp.org/research/systematics-evolution/resources/molecular-biology/facilities/
American Museum of Natural History, Ambrose Monell Collection for Molecular and Microbial Research	The Ambrose Monell Cryo Collection (AMCC) supports a broad range of comparative genetic and genomic research initiatives focusing on earth's biodiversity, including animal, fungal, plant, and microbial diversity. Collecting kits and equipment are also available for sampling and shipping of genetic materials.	http://research.amnh.org/genomics/Facilities/AMCC
Bell Museum of Natural History, University of Minnesota	The Bell Museum of Natural History at the University of Minnesota maintains frozen tissue collections from amphibians, birds, reptiles, and fishes. Tissues are stored at -70°C or in 95% ethanol and are available for research purposes.	http://www.bellmuseum.umn.edu/
Botanical Research Institute of Texas	The Botanical Research Institute of Texas maintains a small repository of genetic material from Cyperaceae. Our current collection includes 233 individual vouchers, with 843 additional samples representing various populations.	http://www.brit.org/

Figure 1. The new web feature compiled and maintained by iDigBio provides researchers with easy access to genetic resources in the United States. As depicted in this screenshot, collections are listed alphabetically and include a link to the full record (see Figure 2), a descriptive abstract, and active URL.



Florida Museum of Natural History, Genetic Resources Repository (University of Florida)

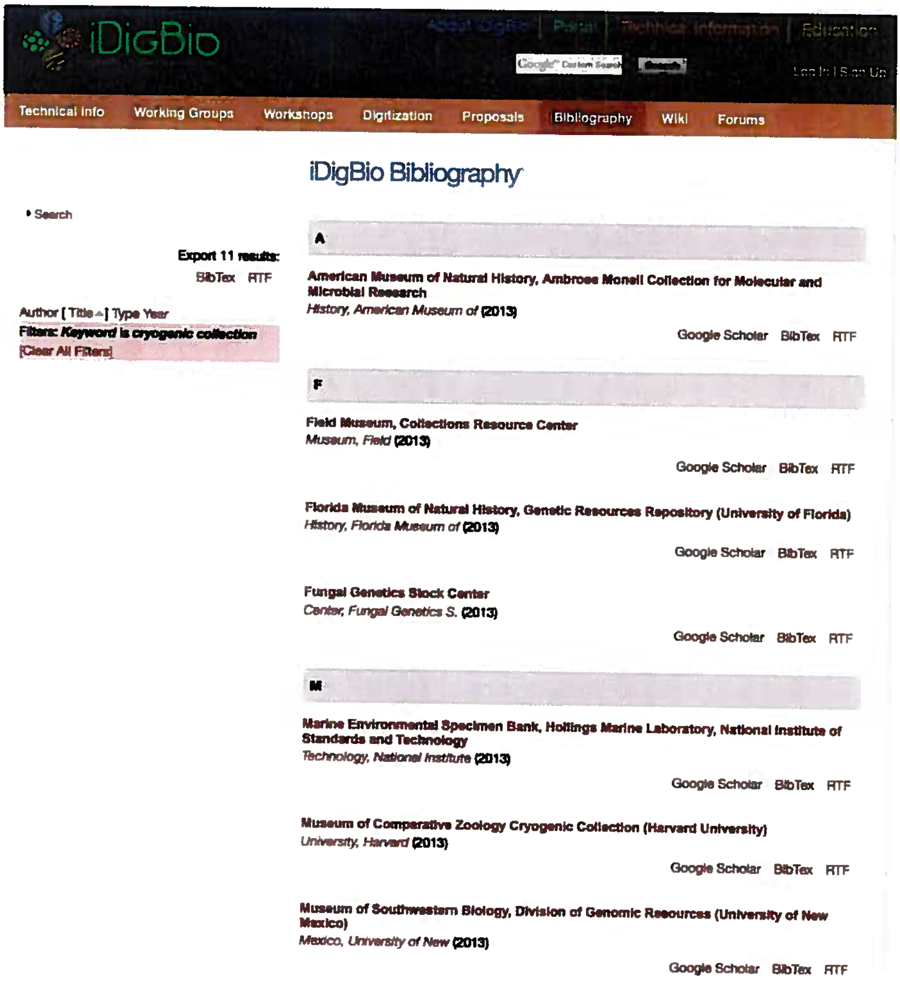
Title	**Florida Museum of Natural History, Genetic Resources Repository (University of Florida)**
Publication Type	**Website**
Year of Publication	**2013**
Authors	History, Florida Museum of
Keywords	-180C, botany, centralized repository, cryogenic collection, DNA bank, entomology, frozen tissue collection, genetic resources, herpetology, ichthyology, invertebrates, liquid nitrogen, mammalogy, nucleic acid extracts, online database, ornithology, vertebrates
Abstract	**The Florida Museum of Natural History's Genetic Resources Repository (GRR) archives more than 30,000 tissue samples and DNA and RNA preparations from specimens housed in the Museum's Mammals, Herpetology, Birds, Invertebrate Zoology and Ichthyology divisions, the McGuire Center for Lepidoptera and Biodiversity, the Herbarium, and the Laboratory of Molecular Systematics and Evolutionary Genetics.**
URL	http://www.flmnh.ufl.edu/grr/

log in or google scholar bibtex rtf **to post comments**

Like iDigBio on Facebook | Follow iDigbio on Twitter
Use of this website is subject to iDigBio's Terms of Use and Service Level Agreement.
This page uses Google Analytics • Google Privacy Policy © Copyright 2011

iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Figure 2. Records for individual collections are maintained in iDigBio as part of a modified bibliographic database. This screenshot of a record for the Florida Museum of Natural History Genetic Resources Repository (University of Florida) illustrates our keyword-based organizational schematic, in which individual keyword links describe the attributes of an individual collection and allow iDigBio users to identify easily other collections with an identical attribute of interest (see Figure 3).



iDigBio Bibliography

Export 11 results:
BibTex RTF

Author [Title ▲] Type Year
Filters: ***Keyword*** **is cryogenic collection**
[Clear All Filters]

A

American Museum of Natural History, Ambrose Monell Collection for Molecular and Microbial Research
History, American Museum of (2013)
Google Scholar BibTex RTF

F

Field Museum, Collections Resource Center
Museum, Field (2013)
Google Scholar BibTex RTF

Florida Museum of Natural History, Genetic Resources Repository (University of Florida)
History, Florida Museum of (2013)
Google Scholar BibTex RTF

Fungal Genetics Stock Center
Center, Fungal Genetics S. (2013)
Google Scholar BibTex RTF

M

Marine Environmental Specimen Bank, Hollings Marine Laboratory, National Institute of Standards and Technology
Technology, National Institute (2013)
Google Scholar BibTex RTF

Museum of Comparative Zoology Cryogenic Collection (Harvard University)
University, Harvard (2013)
Google Scholar BibTex RTF

Museum of Southwestern Biology, Division of Genomic Resources (University of New Mexico)
Mexico, University of New (2013)
Google Scholar BibTex RTF

Figure 3. Selection of a collection record keyword redirects users to an iDigBio Bibliography that includes all records sharing the selected keyword. In this screenshot, all iDigBio records with the keyword attribute, "cryogenic collection", are provided to the user in alphabetical order.

tions simultaneously for a given species and/or type of sample. Currently, the DNA Bank Network and the Arctos Multi-Institution, Multi-Collection Museum Database (http://arctos.database.museum/) both provide web-based search capabilities that facilitate identification of vouchered genetic resources. However, the searchable resources available within each of these networks are currently limited to a small number of collections, and collective lack of knowledge regarding the number and location of additional collections hinders expansion of these resources.

The capability for universal searches for both specimens and genetic resources (as well as their linkage) is coming, but may not be available for some time. Currently, iDigBio's portal contains over 11 million specimen records (as of late February 2014), and natural history collections in the U.S. are making rapid progress with specimen digitization. It may be possible for iDigBio to incorporate genetic resources data in the near future, and allow for linkages between specimen data, images, and genetic resources in a single, inclusive portal. In the interim, the iDigBio resource described here provides information on a diverse set of collections and links directly to genetic resource collections and repositories, enabling rapid searching of multiple collections.

Lastly, we consider the important roles the research community can play in establishing a useful network of genetic resources, especially because networked collections are only as useful as the samples they contain. Individual researchers, collectively, maintain and curate the largest proportion of available genetic resources in the U.S. as part of previous and ongoing research projects. These resources may be of value to current or future generations of researchers, including those in other research areas or disciplines. However, the existence of these isolated resources remains unknown to the larger community. Moreover, these collections risk damage or loss due to a variety of factors (e.g. poor curation, freezer malfunction or power outages, retirement, etc.) and may be best transferred to more permanent storage facilities. Proper archival of genetic resources not only ensures the long-term preservation, curation, and sharing of genetic resources, it also ensures the repeatability of scientific results yielded from these resources. Secondly, the research community can contribute to the banking of new genomic resources for future generations. Some large repositories (e.g. the American Museum of Natural History) are beginning to offer useful collection, processing, and transport guidelines, as well as collecting kits and equipment to aid researchers in banking high-quality genetic resources with broad research applicability. While the benefits for the repository and the research community are obvious, these institutional services and resources may also greatly benefit individual researchers by providing a means for easy and effective collection and archival of research materials in a few steps. Archived material can be requested at any time, eliminating the need for in-lab storage and offering an efficient means for sourcing new material (if needed) for subsequent research projects.

LITERATURE CITED

Adams, R. P. 1988. The development of a DNA Bank for the conservation and utilization of DNA. AIBS Annual Meeting 1988. University of California, Davis. Abstract: American Journal of Botany 75 (6, Part 2): 156.

Adams, R. P. & J. E. Adams (editors). 1992. Conservation of Plant Genes: DNA Banking and In Vitro Biotechnology. Academic Press, San Diego.

Adams, R. P. 1993. DNA Bank-Net: an international network of DNA banks and *in vitro* storage institutions. AgBiotech News Inform. 5: 317N–319N.

Adams, R. P. 1997. Conservation of DNA: DNA banking. Pp. 163–174 *in* J. A. Callow, B. V. Ford-Lloyd &

H. J. Newbury (editors). Biotechnology and Plant Genetic Resources: Conservation and Use. CABI Publishing, Wallingford, U.K.

Adams, R. P., M. Zong & Y. Fei. 1999. Preservation of DNA in plant specimens: inactivation and re-activation of DNases in field specimens. Mol. Ecol. 8: 681–684.

Amato, G. 2013. Collections in the 21st century: An overview of issues in storing samples for genomic analysis with an emphasis on the AMNH collection. Abstract for the U.S. Workshop on DNA Banking, Missouri Botanical Garden, Saint Louis, MO.

Dröge, G., K. Barker & P. K. Flemons. 2013. Darwin Core Archive and the Global Genome Biodiversity Network. Abstract for the Biodiversity Information Standards (TDWG) annual conference, Florence, Italy.

Flemons, P. K., K. Barker & G. Dröge. 2013. Workshop: Darwin Core DNA and tissue data standard for the Global Genome Biodiversity Network. Abstract for the Biodiversity Information Standards (TDWG) annual conference, Florence, Italy.

Mattick, J. S., E. M. Ablett & D. L. Edmonson. 1992. The Gene Library: Preservation and analysis of genetic diversity in Australasia. Pp. 15–35 *in* R. P. Adams R. P. & J. E. Adams (editors). Conservation of Plant Genes: DNA banking and In Vitro Biotechnology. Academic Press, San Diego.

Rice, N., R. Henry & M. Rosetto. 2006. DNA banks: a primary resource for conservation research. Pp. 41–48 *in* M. C. de Vincent & M. S. Andersson (editors). DNA Banks—Providing Novel Options for Genebanks? International Plant Genetics Resources Institute, Rome.

Wong, P. B. Y., E. O. Wiley, W. E. Johnson, O. A. Ryder, S. J. O'Brien, D. Haussler, K.-P. Koepfli, M. L. Houck, P. Perelman, G. Mastromonaco, A. C. Bentley, B. Venkatesh, Y. Zhang, R. W. Murphy & G10KCOS. 2012. Tissue sampling methods and standards for vertebrate genomics. GigaScience 1:8.

Zimkus, B. 2013. Developing best practices for genetic resource collections associated with natural history collections. Abstract for the U.S. Workshop on DNA Banking, Missouri Botanical Garden, Saint Louis, MO.