Culture Collections

National Plant Microbial Germplasm System

A National Initiative to Ensure Essential Resources for Research, Education, and Economic Competitiveness

August 2009

Table of Contents

List of Acronyms and Abbreviated Terms

APHIS – Animal Plant Health Inspection Service

APS – American Phytopathological Society

ARS – Agricultural Research Service

ATCC – American Type Culture Collection

BRC – Biological Research Centers

CBS – Centraalbureau voor Schimmelcultures

CDC – Centers for Disease Control

CiN – Cyber-infrastructure for NPMGS

EBRC – European Biological Research Centers

EBRCN – European Biological Research Center Network

EPPO – European Plant Protection Organization

FGSC – Fungal Genetics Stock Center

FRC – Fusarium Research Center

GBRCN – Global Biological Resource Center Network

GIS – Geographic Information System

GRIN – Germplasm Resource Information Network

ICCS – International Culture Collection System

LSCBR – Living Stock Collections for Biological Research

MTA – Material Transfer Agreement

NCGRP – National Center for Genetic Resource Preservation

NIH – National Institutes of Health

NPGS – National Plant Germplasm System

NPMGS – National Plant Microbial Germplasm System

NSF – National Science Foundation

OECD – Organization for Economic Cooperation and Development

PPB – Public Policy Board

SME – Subject Matter Experts

SOPs – Standard Operating Procedures

USDA – United States Department of Agriculture

USFCC – United States Federation of Culture Collections

WFCC – World Federation of Culture Collections

National Plant Microbial Germplasm System

A National Initiative to Ensure Essential Research Resources

1. Executive Summary: Plant associated microbial germplasm, a resource at risk:

Plant associated microbes play diverse and critical roles in plant ecosystem health, and thus have been a major focus of research resulting in many large geographically-dispersed collections of microbes. These microbial germplasm collections represent an essential foundation for U.S. science to solve a myriad of practical challenges to our agricultural and environmental systems and to advance fundamental understanding of microbial processes. They are irreplaceable and invaluable resources. Microbial collections are at risk however because the United States lacks a coordinated national system to protect, preserve and enhance these valuable resources. There is a critical need to establish a viable and well-coordinated national microbial germplasm system that safeguards the diversity of plant-associated microbial germplasm collections and supports research, education, and practical uses of archived collections. Such a system will catalog and serve as a repository for the historical, genetic and phenotypic variation of plant microbial germplasm. These collections will provide scientists with the resources to solve practical problems ranging from discovering microbes that control plant diseases and weeds, producing valuable pharmaceuticals and industrial enzymes, and to distinguish between a naturally evolved disease-causing organism and an intentionally introduced plant pathogen.

This national initiative outlines a cost-effective model to build a strong National Plant Microbial Germplasm System (NPMGS) composed of distributed expert curated, taxon-specific repositories linked through a searchable common cyberinfrastructure database. A central repository to house backup cultures and receive decommissioned collections will be established. This network of linked collections will ensure that reference strains are no longer lost and that they remain readily available for facilitating comparative research, especially of emerging and reemerging pathogens. This permanent repository system will be accessible by the broader scientific community as well as by law enforcement and homeland security officials responsible for safeguarding our national agricultural system.

The infrastructure envisioned for this national system will be connected to active research programs and their associated expert personnel and will be a joint venture involving a new Federal government initiative, existing structures such as the U.S. National Plant Germplasm System (NPGS), existing university collections maintained in an ad hoc manner as well as individual or collaboratively awarded proposals that may be funded by the USDA, the National Science Foundation (NSF), and other agencies. Since the long-term success of the proposed system will depend heavily on active participation by curators of existing collections and associated research communities, the design and management of the system must balance the responsibility of participation with benefits.

2. Background. Knowing the genotypic and phenotypic variation of plant-associated microbes is essential for providing effective and timely disease management and promoting plant health through research, teaching and extension in the microbiological sciences including plant pathology. In addition to knowledge about the microorganisms themselves, collections provide an essential genetic link between the past and present research and disease outbreaks, and can, for example, explain how changes in newly developed cultivars or plant material influence disease outbreaks. Without access to phenotypic and genotypic diversity associated with past and present pathogen cultures, it is impossible to predict durability of plant resistance. Advances in molecular techniques have revealed that microbial pathogen diversity is much more extensive than suspected previously.

Understanding this diversity is fundamental to systematic cataloguing of fungal, bacterial, nematode and viral strains in order to provide correct identifications for regulatory compliance or attribution for forensic purposes. Reference cultures with phenotypic and genotypic fingerprints and geospatial and temporal contexts facilitate identification of emerging pathogens and monitoring of pathogen migration and evolution.

2.1. Value of plant associated microbial germplasm to the nation.

2.1.1 Advancement and competitiveness of science*:* It often takes years to establish the groundwork necessary to study economically important microbial groups. Loss of isolates previously studied often prevents other scientists from confirming and building upon past studies. Culture collections must be connected intimately to active research programs and their associated expert personnel and data resources in order to be increasingly relevant and useful. Realizing the potential of microbial genomics as a foundation for developing effective disease control and biological discovery also hinges on how effectively we use the sequenced isolates as a reference to understand the genetic and phenotypic diversity within a pathogen species.

Genomics is more than studying gene function on a large scale. It should also serve as a foundation for understanding the genetic basis of phenotypic diversity and variation within species as well as species evolution. Genome sequencing and subsequent functional analyses typically concentrate on a single or small number of isolates in target species. However, knowing the complete genetic makeup of a few isolates is not sufficient for understanding fully the genotypic and phenotypic diversity within and among species. For instance, certain microbial pathogenicity genes are present on a 'dispensable' chromosome and, consequently, are absent in some isolates of the same species. Therefore, genotypes and phenotypes of diverse isolates need to be characterized using the sequenced strain as a reference. As new high-throughput and sequencing technologies become widely available and economical, re-sequencing of multiple isolates as a means for comparing genome-wide differences is becoming routine and cost efficient. The motivation for linking genome sequences to diverse isolates within sequenced species is to realize the full potential of genomics as a foundation for understanding the genotypic and phenotypic diversity within and between species and to guide re-sequencing.

2.1.2 Resource for controlling problems and harnessing the benefits from microbes*:* In effect, microbial culture collections are libraries of the genotypic and phenotypic diversity of microbial communities and their spatial and temporal structures. The microbial germplasm collections help us determine the identity of new microbial isolates and trace the diversity and movement of microbes in space and time. This capability is particularly important to protect the agricultural and environmental systems in the nation from major diseases. The introduction of exotic pathogens and variants of endogenous pathogens have accelerated greatly as a result of the globalization of commerce, with its accompanying movement of more materials through multiple entry points and increased human travel. *Phytophthora ramorum*, a recently discovered pathogen, has caused sudden oak death in west coast forests and diseases in ornamental plants throughout both the United States and Europe. Although it has been several years since *P. ramorum* was first identified, its origin remains unknown. This detective work would have been much easier with comprehensive data on the global diversity of *Phytophthora*. As exemplified by the recent discovery of *P. kernoviae* on rhododendron, beech and oak in the U.K. and custard apple in New Zealand (www.eppo.org) where *P. kernoviae* was confirmed to be present for decades based on culture collections data, *P. ramorum* is not the last *Phytophthora* threat to the U.S. Considering the vast magnitude of unexplored microbial diversity, novel plant pathogens will continuously be identified.

After the September 11 attacks and the subsequent anthrax releases, it became abundantly clear that the threat to agriculture from the deliberate release of pathogens cannot be overlooked. Enhancing our capability of rapid disease detection and diagnosis will significantly increase the probability of achieving containment and eradication of high-risk pathogens. Timely identification,

risk assessment, and monitoring of pathogen movement depend on rapid access to cultures that represent the known diversity of plant pathogens along with their associated genotypic, phenotypic, and epidemiological data. Archiving pathogen cultures and associated data to support such activities should be an important step in enhancing nation-wide preparedness against plant biosecurity threats.

Preservation of pathogen isolates from past disease epidemics is equivalent to archiving key documents needed for understanding an important historical event. As Santayana said, “those who cannot remember the past are condemned to repeat it," our failure to learn from past and present disease epidemics will result in the likelihood of repeatedly suffering from the same epidemic scenario.

As recyclers of organic matter and as symbionts of many terrestrial plants, microbes are essential components of healthy ecosystems. For thousands of years, humans have used microbes for the production of foods, including the use of fungi, yeasts and bacteria to produce bread and fermented foods and beverages. Certain species have been invaluable for the discovery, development and manufacture of useful compounds, including pharmaceuticals, organic acids, industrial enzymes and recombinant proteins. Considering the diverse metabolic capabilities in the limited number of commercially utilized species, microbes represent a vastly under-utilized bioprospecting resource.

2.1.3 Foundation supporting the exploration and monitoring of the global microbial diversity*:* The under-explored microbial diversity in nature underscores the importance of cataloguing microbial germplasm in a format that supports future exploration. Although plant associated microbes have been a major focus of research due to their essential roles in plant health, we still remain overwhelmed by their diversity, most of which is unstudied. The Kingdom Fungi is estimated to consist of approximately 1.5 million species, yet we probably know less than 100,000 of them. Application of DNA sequence-based identification techniques further underscores how little we understand about the microbial diversity in nature; as much as 99% of bacteria from many environments cannot be isolated using standard culturing techniques. Indeed, many bacterial species are known to exist only from gene sequences from environmental samples. Considering that we have barely begun to survey microbial communities in nature using molecular techniques, we will continue to uncover many previously undiscovered microbes as we explore uncharted environments.

2.1.4 Missed opportunities due to the lack of appropriate microbial germplasm*:* It is argued that microorganisms can always be collected again and will not be ‘lost’ forever. However, science has shown that rare and unique strains exist and may never be re-isolated. These unique strains provide value in forensics, plant breeding, commerce and biological control of pathogens and pests, as well as for the continuity of science, which replicates and builds upon previous findings.

Despite the importance of microbial germplasm reference material, the survival of this cache of genetic diversity is threatened because most archived cultures in the United States exist only in individual scientists’ collections at academic institutions without support for long-term curation. Once the researcher retires or leaves the institution, the microbial germplasm frequently is lost even when paid for by public funds. Because a centralized database of microbial germplasm does not exist, these individual collections are not uniformly catalogued and available to the broad research community.

Valuable collections are maintained in the private sector. At present, research is impeded and often duplicated due to the inaccessibility and attrition of historical samples. While most industry collections are considered proprietary, a company may be willing to offer some or all of their collections to a national system, similar to what has occurred with plant germplasm and genomic resources.

An example of a missed opportunity because an important collection was lost relates to *Fusarium* head blight of wheat which has reemerged as a severe fungal disease on wheat and corn throughout the grain producing states and Canada. The renowned scientist, W.L. Gordon conducted the first high quality surveys of *Fusarium* species in the U.S. and Canada in the 1940s and 1950s.

His culture collections were discarded following his retirement because of no succession plan to preserve and maintain his fungal collections. This collection would have been a critical resource in understanding and solving what is now one of the most important challenges to agriculture in North America.

Finally, because there is no common database of materials, it is not possible to determine the extent of plant microbial germplasm available for use by the broader scientific research community. Existing collections, such as the American Type Culture Collection (ATCC), are limited in scope and not characterized extensively beyond information provided in the initial submission and in many cases strains are cost prohibitive to obtain. It is clear that such collections cannot substitute for the network proposed here.

2.1.5 Collections at Risk. At least 10 major collections of rare or non-duplicated strains are at risk of being destroyed due to budget cuts and retirements, especially at universities (see Appendix 1). For example, a university research scientist who has 2,552 pathogenic and beneficial strains of bacteria, including possibly the largest collection in the United States of Gram-positive bacteria, as well as unique strains, expects to retire in early 2010. No one at the home institution is willing to be the recipient of and curator for this collection. Also, the World Phytophthora Collection at the University of California at Riverside annually is at risk due to lack of funds for maintaining this world-class collection.

2.1.6 Success stories resulting from the proactive characterization of culture collections. Over the course of several years, scientists at the Fusarium Research Center (FRC) at Penn State and the Microbial Genomics and Bioprocessing unit at the U.S. Department of Agriculture (USDA) facility in Peoria, Illinois collaborated to characterize genetically all *Fusarium* cultures in these two collections that were associated with human infections. *Fusarium* infections of humans, which include deadly, invasive infections of severely immune-compromised individuals as well as serious infections of the cornea, are often unresponsive to antifungal drug therapies. As these studies neared completion in early 2006, news broke of an outbreak of corneal infections in Southeast Asia and the United States associated with the use of a particular contact lens solution. Because clinical *Fusarium isolates* were characterized proactively, FRC and USDA scientists were able to provide the Centers for Disease Control (CDC) and the manufacturer with precise, DNA-based identification of the fungi associated with the infections, and their genetic relationships with the previously characterized species, in a matter of a few days. As a result, investigators were able to quickly determine that there were multiple species of *Fusarium* associated with the infections, and that they were most likely introduced from the patient environment and not from contaminated product. This information was key in the development of an infection model for this disease, and the discontinuation of the product. If clinical isolates from these collections had not been proactively studied, these investigations would have taken many weeks or months to complete rather than a few days.

Another positive example in the U.S. is the situation with os-2 mutants of *Neurospora crassa.* This mutant was available in the FGSC collection for many years but was rarely utilized other than as a genetic marker for studying unrelated phenomena. In 1999, however, it was shown that mutants in this gene also conferred resistance to phenylpyrrole fungicides. Orders for strains carrying this mutation jumped over 300 percent and subsequent work identified the basis of resistance.

2.2 Expected impacts from the NPMGS

2.2.1 Readily available resource supporting industrial bioprospecting: Opportunities are expanding for use of microbial strains as industrial catalysts. Enzyme usage is increasing as an environmentally friendly alternative technology. Reduced usage of chlorinated chemicals in the pulp and paper industries and associated lowered costs are now a reality. In the textile industry, improved products for softness, smooth finish, stone-washing and reduced pilling and fuzz are available through microbial processes. Bioethanol production from waste is a growing industry related to reducing high fossil fuel demands. Many novel compounds with useful properties remain to be discovered from various microbes. Given the increasing demand for alternatives to chemical control for plant diseases, exploration of diverse microbes as biocontrol agents to protect plants from biotic and abiotic stresses will increase.

2.2.2 Readily available resource supporting scientific inquiry and education: Advances in molecular biology, microbial genomics, and biotechnology in the past decade have fundamentally changed the nature of research on plant associated microbes, making it possible to conduct many types of experiments and analyses using molecular database information, or microbial nucleic acids, in addition to living cultures. Such advances are useful in studying all microbes, but have impacted especially work on unculturable organisms and microbial communities (metagenomes). Where maintenance of living cultures is not practical, or is prohibitively expensive, libraries of molecular components (DNA, RNA, proteins, etc), while not a replacement for live culture collections, are becoming central to scientific inquiry. These resources also play an essential role in educating the next generation of scientists since they support students to learn how to study the diversity, evolutionary history and functional diversification/innovation of microbes important for the nation's agriculture, food, and environmental systems, at levels ranging from molecules to ecosystems.

2.2.3 ***Forensic support for the identification and monitoring of major pathogens:*** The importance of culture collections was exemplified by the intentional dissemination of *B. anthracis* through the U.S. postal system in 2001. A national microbial culture collection could expedite the analysis and characterization of biological evidence to identify a pathogen's genotype, its genetic uniqueness, or the genetic variations within populations of similar isolates. The ability of the Federal Bureau of Investigation to investigate a terrorist attack with an agricultural pathogen would be significantly enhanced by the establishment of a national microbial culture collection. Another example of the value of collections to human health as discussed above, was the rapid and accurate identification of contaminants in a contact lens solution made possible because of the *Fusarium* research center culture collection and the associated data (e.g. sequences, morphology, biology, distribution) associated with the cultures ("value added cultures").

2.2.4 Enhanced global cooperation: Many threats to the agricultural and environmental systems in the nation are global in nature and should be addressed through close cooperation with other nations. The NPMGS will allow us to network closely with the international microbial germplasm collections such as the Canadian culture network, national (U.S. Federation of Culture Collections), international (World Federation of Culture Collections), and commercial resources. National collections in other countries such as Belgium and the Netherlands, Centraalbureau voor Schimmelcultures (CBS), are more robust, and the U.S. stands to benefit from such a coordinated effort. Management of the NPMGS should be compatible with the OECD and its plan for establishing a Global Biological Resource Center Network (GBRCN). Leveraging of funds to support the NPMGS may be possible from these associations.

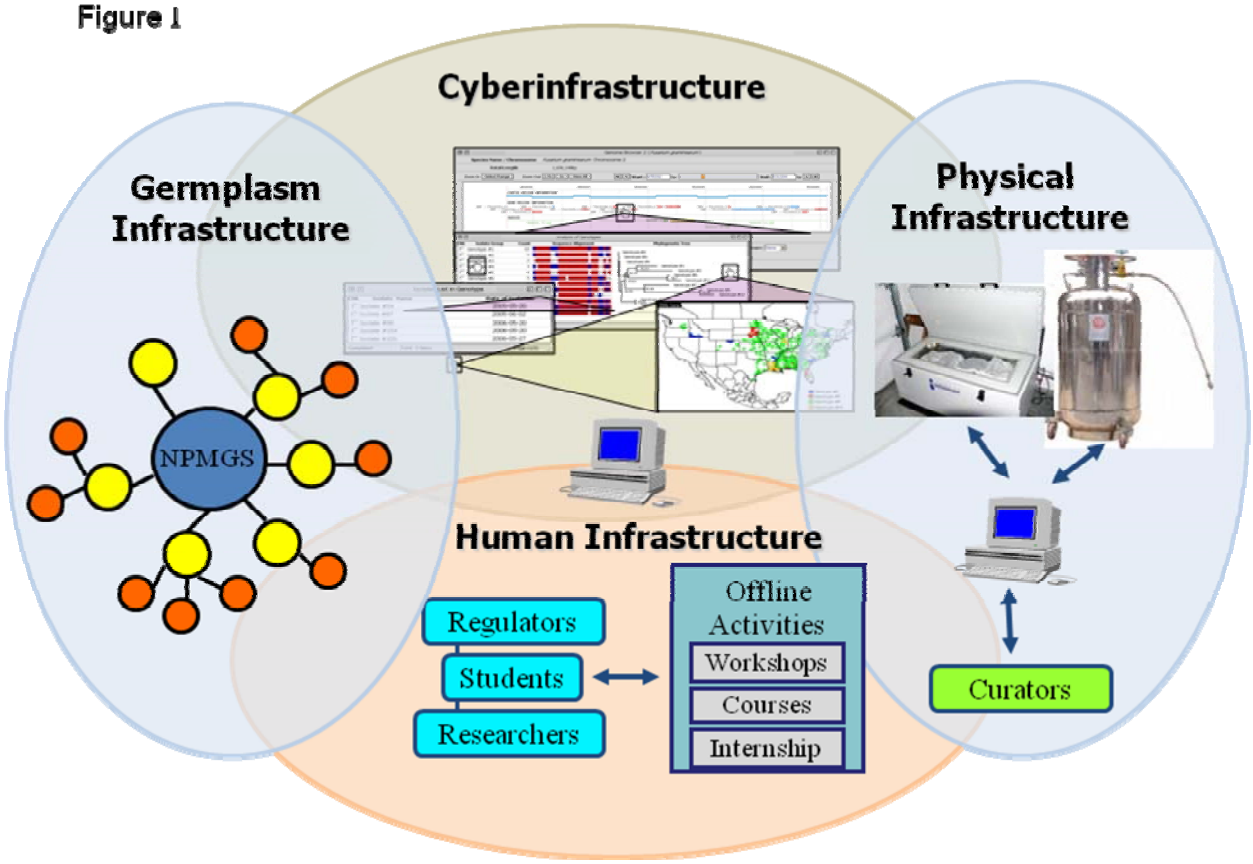
Beginning in 2000, the Organization for Economic Cooperation and Development (OECD) with 30 member countries, established a Task Force on Biological Resource Centers in partnership with a European Union project that founded the European Biological Center Network (EBRCN) linking

major European Biological Research Centers. The mission of this Task Force is to extend and develop, on a global scale, the best practice and credibility standards for Biological Resource Centers that were established earlier. A Global Biological Resource Center Network (GBRCN) is envisioned that will coordinate and raise the standards of major Biological Resource Centers worldwide *(OECD Best Practice Guidelines for Biological Resource Centres) DSTI/.STP/BIO (2007)9*. The Task Force has put forward several key documents such as *Biological Resource Centers: Underpinning the future of life sciences and biotechnology (2001); Review of the current status, activities and future of existing Biological Resource Centers (2001); and Guidance for the operation of Biological Research Centers (BRCs) (2004).* While there has been some American representation in the planning of the GBRCN, coupled with the NPMGS, this global network is a tremendous opportunity to establish international standards for research centers maintaining microbial cultures.

3. NPMGS Overview. We propose a new paradigm for plant microbial resources in the United States. Specifically, we propose the establishment of a National Plant Microbial Germplasm System (NPMGS) of plant-associated microbial resources composed of multiple elements, including "traditional" living culture collections, libraries of molecular components, and a centralized searchable database with supporting information technology (IT) tools. A database with strain information from each specialized center will link the separate repositories and strains deposited in a back-up facility. This system will ensure the preservation and safeguarding of collections of living plant-associated microorganisms that represent the broadest range of phenotypic and genotypic diversity in a permanent repository system together with a permanent, robust database that provides the broader scientific community ready access to critical information.

***3.1 NPMGS Management and Structure*:** The NPMGS will consist of four main elements: (i) the germplasm collection and various types of phenotypic and genotypic data associated with the germplasm at individual Biobank Centers ("Germplasm infrastructure"), (ii) communication mechanisms and support that create well-connected networks of researchers and regulators whose work is founded on the cultures at individual Biobank Centers ("Human infrastructure"), (iii) a physical and logistical support system to better preserve and manage the microbial germplasm at individual Biobank Centers ("Physical infrastructure"), and (iv) informatics platforms and tools that support and link the other three components ("Cyber-infrastructure") (Figure 1).

Figure 1



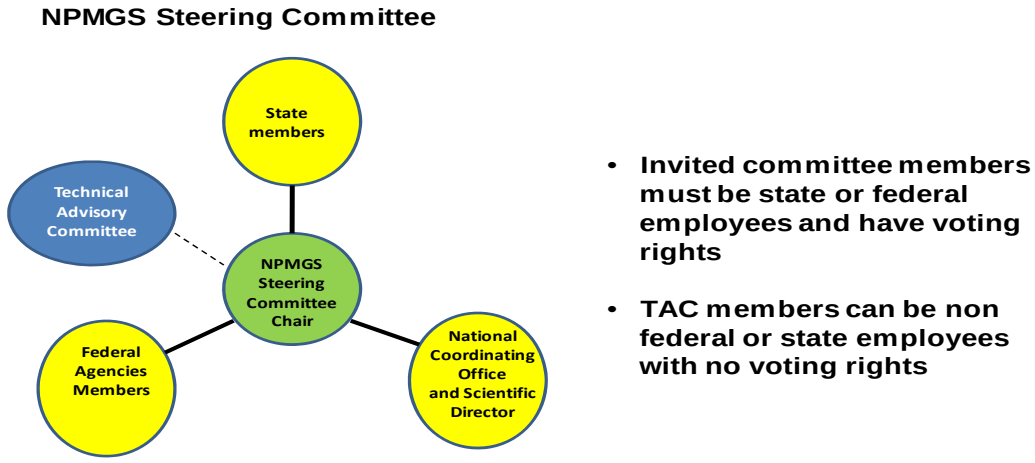
3.1.1 Germplasm infrastructure: Taking advantage of existing expertise and collections, a stably funded, interconnected system of specialized centers ("Biobanks") together with a central backup facility is proposed for organizational coordination and management, and for the physical maintenance and distribution of strains. Each Biobank Center in the system will be organized around groups of organisms or areas of plant health. Varying in size and organismal focus, the Centers will maintain, authenticate and may distribute specific groups of organisms. Each group of microorganisms presents special challenges for maintenance and characterization of strains. Some collections are specialized in pathogens of specific crops e.g., wheat and soybeans and these may overlap with other microbe specific collections of viruses, bacteria, fungi, or nematodes. Having some overlap ensures that resources would be available to accept endangered or orphaned culture collections from retiring scientists or companies and will serve as a resource for those working in that area of research.

The value of culture collections depends on data (actual or potential) associated with the cultures, including basic accession information (geographic origin, host or substrate of origin, equivalent accessions in other collections, species name) and research data generated from that culture. This may include nucleotide sequence information (DNA barcodes, fingerprints, and genetic markers), secondary metabolite profiles, disease syndromes (symptoms, hosts), associated literature and web links, and other relevant information. To realize fully the potential of culture collections for addressing the needs of research, regulation and response in plant health, these data must be well curated, readily available, and distributed with the cultures. Considering that most of the materials and data typically will come from individual researchers, it is essential to establish mechanisms that encourage and support the deposition of relevant materials and data by members of the research and regulatory communities.

3.1.2. Human infrastructure: The NPMGS should be guided by a steering committee composed of Federal employees who will provide oversight for the development and continuing operation of the NPMGS (Figure 2).

Figure 2

NPMGS Management



This steering committee will define a model to build a strong NPMGS for the U.S. by taking advantage of the existing U.S. wide National Plant Germplasm System (NPGS) and the Germplasm Resource Information Network (GRIN). The proposed NPMGS will consist of specialized collection Biobank Centers associated with experts as curators and each directed by a Biobank Manager.

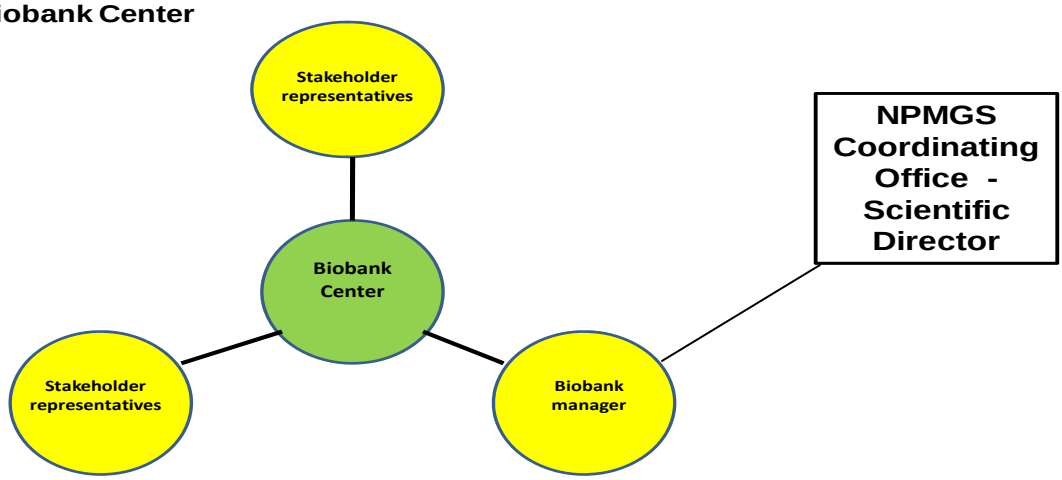
A governance structure for the steering committee is proposed based mainly on Canadian and European models but modified to fit the U.S. situation. We envision establishment of an NPMGS coordinating office staffed by a scientific director who will interface directly with the steering committee regarding the activities conducted at each Biobank Center. Input regarding the operation of each Center will be provided by a Biobank Center manager directly to the scientific director who will be located at the NPMGS centralized coordinating office (Figure 3). Each Biobank manager should participate in a national or international committee that reviews the activities of their collection. These may be formal advisory boards (like that supporting the Fungal Genetics Stock Center), or created *ad hoc*. The scientific director will also be a non-voting member of the steering committee.

We also propose the establishment of a Technical Advisory Committee (TAC) that would interface with stakeholder groups and advocates of the NPMGS for public support and accountability and provide suggestions and recommendations to the Steering Committee (Figure 2). Membership in this advisory committee will be recommended by relevant stakeholders and professional societies, such as the American Phytopathological Society (APS), and appointed by a designated USDA representative. Advisory committee members will be non-Federal stakeholders and advocates of the NPMGS with knowledge of broad policy issues related to agricultural science. They will provide recommendations to the steering committee on policy and budgetary issues. In addition to the Advisory Committee, we propose the development of other mechanisms to afford the greatest opportunity for input from a broad range of stakeholders, including, for example, interactive websites and stakeholder listening sessions.

Figure 3

NPMGS Management

Biobank Center



The NPMGS coordinating office staffed by the scientific director will be located at the National Center for Genetic Resource Preservation (NCGRP) in Ft. Collins, Colorado, where USDA, Agricultural Research Service (ARS) currently maintains back-up germplasm of plant accessions and have adequate facilities and equipment in place. The scientific director and steering committee will seek to balance the responsibility of each specialized Biobank Center with expected benefits. Activities, as defined by the steering committee, will also be coordinated with existing collections, such as the American Type Culture Collection (ATCC) and international collections in Canada, Europe, Australia and others. Collections at ATCC and others may be limited in scope and not extensively characterized but efforts will be made to provide linkage to them.

Membership to the steering committee will be made by appointment by a designated USDA representative for a five year term. One member of the steering committee will rotate off the committee each successive year and be replaced by a new incoming member. This rotation ensures continuity on the committee while providing balance. The steering committee will develop criteria for including new collections and assessing the condition of collections that will be included in the NPMGS. These criteria will also address what actions are necessary if expertise were discontinued at a specialized Biobank Center (i.e. curator retires) and whether collections could be folded into other existing collections. The steering committee will meet on a regular basis to identify additional microbial strains and isolates to be included in the national collection based on these criteria. The NPMGS steering committee will be responsible for identifying orphaned collections and collections at risk and for specifying what actions are needed to save endangered culture collections. It is recommended that the steering committee should be responsible for developing criteria to prioritize endangered and threatened collections and to determine where they will be preserved.

An essential component to enhance the materials and data associated with the NPMGS and translate the knowledge derived from them into improved understanding of microbes and practical problem solving is a network of researchers, educators, and regulators who work closely together, share knowledge, and help train the next generation of scientists and educators. As the complexity of biological inquiry increases, the importance of and need for team-based problem solving also increases. With 'specialists' of different expertise working closely together across institutional and geopolitical boundaries, a mechanism is needed that will facilitate and support such cooperation. The NPMGS will facilitate and support team building for the Biobank Centers by providing tools and data via the web.

Given the vast diversity of microbes in relation to the dwindling number of trained field scientists and systematic biologists, a long-term investment in related training is essential to ensure the effectiveness of the NPMGS. In a 2004 meeting, the European Plant Protection Organization (EPPO) declared a state of emergency in plant health, stating that "taxonomy, classical plant pathology and other scientific fields which are vital for sustaining sound public policy are threatened with extinction, because they are no longer in the forefront of science priorities," and recommended urgent action to prevent the disappearance of these disciplines. Recent APS surveys (*Phytopathology News,* Vol. 42, No. 1) and an APS sponsored workshop (http://www.apsnet.org/online/proceedings/EducationWorkshop/) revealed a lack of applied training of graduate students in all core areas of plant pathology and a lack of skilled and broadly trained plant problem diagnosticians. Without adequate human capital, the future of the plant associated microbial germplasm collections and our preparedness for agricultural and environmental safety and security will be jeopardized. The steering committee and the advisory committee, in collaboration with stakeholders, will also develop action plans to support and enhance the established scientific community and to ensure the education of future scientists.

3.1.2.1 Culture Distribution and Cost Recovery: Microbial cultures should be distributed to all legitimate users throughout the world regardless of whether they are government, non-profit or commercial users. It will be incumbent upon the steering committee to make a final determination in cases in which there may be concern over the legitimacy of the requesting party or country of origin when State Department restrictions apply. Otherwise, the Biobank manager at the specialized Biobank Center housing the collection will decide if requests will be honored. However, all users who request plant pathogenic strains are required to have a valid USDA Animal Plant Health Inspection Service (APHIS) permit to receive those strains. A nominal fee may be charged to requesters to discourage frivolous strain requests and to defer shipping costs. The steering committee will determine appropriate fees and review special circumstances such as when collections will be used for educational purposes and for research directly involving curators.

3.1.2.2 Material Transfer Policies: Because every transfer of materials has explicit or implied expectations, the NPMGS steering committee will examine the requirements for a Material Transfer Agreement (MTA) with input from relevant advisory committees and in cooperation with the USDA Office of Technology Transfer (OTT). Resources relevant to the NPMGS MTA include the USDA-OTT and National Institutes of Health (NIH) confidentiality agreements and their MTAs which can be found at: http://www.ars.usda.gov/Business/docs.htm?docid=2075 and http://www3.niaid.nih.gov/about/organization/odoffices/omo/otd/pdf/UBMTA.htm respectively. The NPMGS MTAs will allow distribution of cultures to any non-profit recipient engaged in research with biological materials. It will stipulate that the recipient may not re-distribute original materials but may distribute modified materials. The Biobank Centers of the NPMGS are not expected to be International Depository Authorities according to the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure (1977). Materials for deposit in the back up collection at the National Center for Genetic Resources Preservation in Fort Collins, CO. will be accepted under a modified transfer agreement that mandates a five year re-evaluation to assure that the depositing collection is still responsible for the materials.

3.1.3 Physical infrastructure: The germplasm infrastructure envisioned will be connected to active research programs and their associated expert personnel and will be a joint venture involving a new Federal government initiative, existing structures such as the U.S. wide NPGS, existing university collections maintained in an ad hoc manner, as well as individual or collaborative, competitively awarded proposals that may be funded by the USDA, the NSF, and other agencies. Since the long-term success of the proposed system would depend heavily on active participation by existing collections and associated research communities, the design and management of the system must balance the responsibility of participation with benefits.

The system is envisioned to include 10-30 specialized Biobank Centers covering a wide range of plant-associated microorganisms. Each center would require a stable source of funding, the amount of which would depend on the size of the collection, complexity of strain maintenance, identification and characterization methodologies as well as distribution request volume. The bulk of the funds for NPMGS would support the specialized Biobanks including equipment and associated clerical, technical and research personnel.

Each Biobank Center may be connected with one or more satellite Centers, and be managed by a group of experts (see section 3.1.2). A database with strain information from each Biobank would link the separate repositories and all strains will be deposited in a back-up facility. This system will ensure the preservation and safeguarding of collections of living plant associated microbes that represent the broadest range of phenotypic and genotypic diversity in a permanent repository system together with a permanent, robust database that provides the broader scientific community ready access to critical information.

Individual Biobank Centers must be managed according to standard operating procedures (SOPs) for culture authentication, preservation, distribution, and documentation. The establishment of SOPs will be one of the first actions by the NPMGS Steering Committee. These SOPs may be similar to the World Federation for Culture Collection “Guidelines for the Establishment and Operation of Collections of Cultures of Microorganisms” 2nd Edition, as published June 1999. Support for migrating “collection-specific” databases into the NPMGS database system and for maintenance and improvement may be one means of inducing independent collections to participate in the NPMGS. Support for the NPMGS would encompass the development of software platforms and tools that support the data archiving and management needs of the participating Biobank Centers and research communities that use the archived cultures.

Most strains should have duplicates deposited at a centralized location so that, if the strain is lost from the specialized center, it can be retrieved from the back-up location. In one scenario, the back-up facility could be the point of distribution of strains. Some strains such as those representing a population genetic survey might not have every isolate backed-up, but most unique strains would be backed up at the centralized location.

Appendix 1 lists some existing specialized centers and the estimated number of strains currently held. Criteria must be developed to determine which collections should serve as NPMGS sanctioned Biobanks. To qualify for funding, each Biobank Center would need to maintain a specified level of accountability. Simply housing the strains would likely not be sufficient. Taxonomically accurate and up-to-date nomenclature should be maintained and strains distributed to legitimate users upon request by NPMGS repositories.

The USDA ARS currently maintains back-up germplasm of all plant accessions at the National Center for Genetic Resource Preservation (NCGRP) in Fort Collins, Colorado. This facility has the space and most of the equipment needed for the back-up storage of strains in the NPMGS. Most, but not all, plant-associated microorganisms can be lyophilized and/or cryopreserved or stored as glycerol stocks at –80 C in small vials and thus take up relatively little space. Funds would be required for the maintenance of this space and equipment as well as the personnel to take care of the acquisition and distribution of back-up strains. The specialized centers periodically would need to send newly acquired strains for back-up at this facility. Following the start-up phase it is anticipated that the back-up facility would have relatively little activity except to maintain and distribute strains as needed. There would be anticipated recurring costs associated with technical personnel, overhead and the occasional need to upgrade equipment.

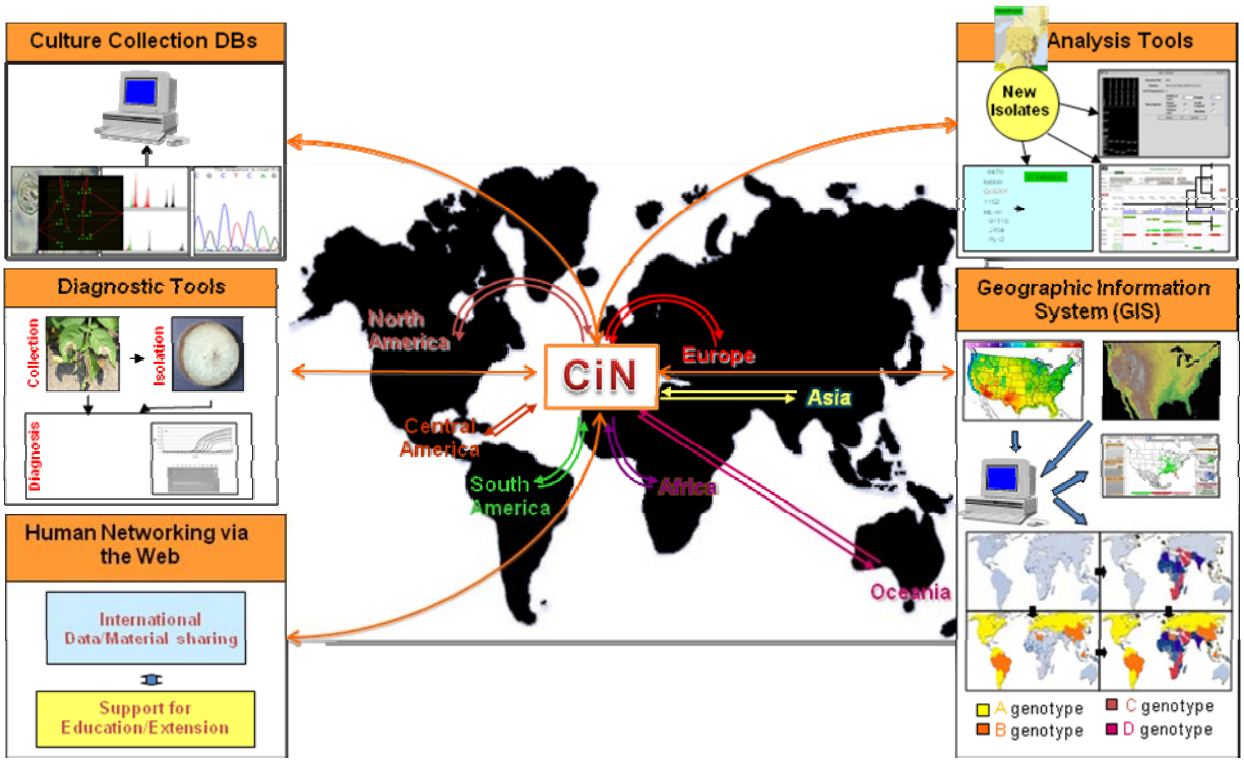
When participating collections join the NPMGS, they must estimate the number of cultures to be stored at the back-up NCGRP and provide a justification and rationale. One potential positive side-effect of NPMGS is the opportunity for USDA-APHIS to more easily manage permits for the movement of plant pests and pathogens. The NPMGS governing body will interact with regulatory personnel at APHIS to develop plant quarantine and other permit regulations that will work in conjunction with the NPGMS. Ideally, an APHIS regulatory administration will serve on the governing body (see section 3.1.2).

3.1.4. Cyber-infrastructure for NPMGS (CiN): Rapid accumulation of data from disparate, yet complementary, areas of research on microbes, such as genomics, molecular genetics, taxonomy, population genetics, and ecology, present both opportunities and challenges to researchers and regulators. Integration of these data can significantly enhance understanding of microbial biology at levels ranging from molecular processes in cells to ecosystem services and will also open new avenues for controlling the problems and harnessing the benefits from microbes. However, the rapid increase in data content and complexity also poses many challenges in transforming data into knowledge. In this age of information overload, a problem we frequently encounter is not the lack of data, but the difficulty of integrating complex and rapidly accumulating data to build a global picture. We cannot overemphasize the importance of archiving available microbe-associated data and knowledge in a format that can maximally support present and future research; failure to address this need will make us 'data rich but knowledge poor.'

To address the increasing need for informatics support in managing and utilizing the NPMGS, we propose a comprehensive cyber-infrastructure, named the **CiN** for **C**yber-**i**nfrastructure for **N**PMGS (Figure 4). The **CiN,** as envisioned, goes well beyond cataloguing and managing the digitized index cards of accessioned cultures. It will support and integrate research, regulatory, education and outreach activities at individual Biobanks and associated communities and will also serve as the hub linking with parallel activities throughout the world. A key component of the **CiN** is a federation

of Biobank-specific databases. All cultures submitted to the NPMGS for backup storage must be accessioned into a searchable master database in the **CiN.** This database should include extant data associated with the cultures in their home collections, including equivalent accession numbers, geographic origin, host, and original depositor. Other fields should include scientific name, pathogenicity, year of isolation, storage location and conditions, location of voucher specimen or source of specimen, type culture, if applicable, date of back-up deposition; strain distribution data, catalogued marker genes analyzed, and other data as needed. As the strain is characterized further, information would be added.

Figure 4



Besides the germplasm-associated data, the **CiN** will systematically archive and link datasets from multiple areas, ranging from genomics data to taxonomy to population and ecological contexts of individual microbial communities, and provide multiple tools to support data management, analysis and visualization. This high-level integration of tools and data will facilitate data-driven modeling and hypothesis development, response to emerging pathogen threats, and efficient transformation of the data into knowledge and also will support education and outreach. For example, the **CiN** will address the growing need for integration of geospatially and temporally referenced biological and environmental data to understand microbial ecology and population biology. This need is particularly acute in managing plant disease outbreaks, because they result from complex interactions among hosts, pathogens, and the environment (a.k.a., the disease triangle). Identifying the cause(s) of disease outbreaks, forecasting their spread, and implementing appropriate disease management strategies require integrative analysis of such disparate data sets. The **CiN** will provide GIS (Geographic Information System) tools to support the visualization of disparate data in spatial and temporal scales and also to link isolate collection with the visualization of collected isolates. A global monitoring system equipped with this proposed functionality will help CiN users better understand microbial ecology and population biology.

These databases may also include associated molecular analysis (genomics, phylogenetics) and GIS tools that can be consulted on line to enable researchers to secure information about isolates of interest and visualize the genetic and geospatial relationships among chosen isolates. At this time, several current IT platforms (e.g., BIOLOMICS, the *Phytophthora* Database, the *Fusarium* Database) exist that may satisfy these needs and may serve as starting points for designing the **CiN.** Biobank-specific databases should be synchronously linked to the core database so that as new data (not necessarily all types of data) become available at a Biobank Center, the content of the core database also becomes automatically updated.

Bioinformatics in **CiN** and all database information will be integrated in the existing U.S. wide, National Plant Germplasm System (NPGS) that is used to manage the 500,000+ accessions of plant germplasm distributed at the plant germplasm centers throughout the U.S. This USDA maintained database, entitled the Germplasm Resources Information Network (GRIN), includes all of the data about the acquisition and existence of plant germplasm as well as their characteristics. The NPGS GRIN system could be expanded to include information about plant-associated microorganisms. Navigation within the GRIN system is primarily through the scientific names of plants and a full time nomenclature scientist continuously reviews and updates the accuracy of the plant names including synonyms. If the GRIN database is used for the NPMGS, it will be essential to fund one or more curators who can function as nomenclature scientists to update the accuracy of the scientific names of plant-associated microorganisms. The ability to locate accurately strains of specific species of interest depends on the continuous updating of scientific names. This will be especially challenging for fungi and bacteria for which there are often two or more scientific names for the same species.

The GRIN database has many of the fields required by the NPMGS but this system would have to be modified to include a number of additional fields that do not exist currently, for example, plant host, symptoms, toxin production and epidemiological data. Start-up funds would be needed for adapting the GRIN database to include plant-associated microorganisms (thus becoming the plant microbial information database portion of GRIN) and for modifying and importing existing data from established culture collections. The estimated initial number of strain records of plant associated microorganisms for the NPMGS is about 200,000 with the expectation that this will grow considerably over time.

Once the NPMGS data are established in **CiN** and integrated with the GRIN system, it will be useful to have the data included in other systems such as *www.straininfo.net*, which includes data from a number of international culture collections. This would also provide a direct route towards linking with an International Culture Collection System (ICCS). Following the start-up phase it is anticipated that each specialized center would enter their own data and thus costs for maintaining the GRIN microorganism data would be relatively stable.

This centralized database must be developed in the short-term, and be online as cultures are deposited in a core facility. Support will be provided to individual Biobank Centers to implement a database system that will include all relevant data associated with cultures. Another major component of the **CiN** is the web-based tools supporting the human infrastructure associated with the NPMGS. Web-based models and tools can potentially bring about major changes in how we conduct research, education and extension/outreach, transform the knowledge and work ecosystems supporting microbial science, and strengthen research communities associated with individual Centers.

4. Costs of Developing and Maintaining a National Plant Microbial Germplasm System: The costs for establishing a system of repositories (i.e. Biobank Centers) with a centralized database and back-up repository is roughly established in Appendix 2. Such a system could be established in phases with emphasis initially on certain endangered groups of organisms or culture collections (see section 2.1.5). The role of culture collections is diverse and critical to national needs involving disease causing agents, beneficial and ecologically vital microorganisms, industrially valuable materials, and biosecurity. The collections acquire, preserve (maintain), identify or verify acquisitions, distribute materials, educate and train future scientists, and analyze or conduct research on collected materials. All these functions require monetary support.

***4.1 Central Hub, CiN Development*:** Modify and populate GRIN system. Start-up costs used to modify GRIN (USDA-ARS Germplasm Resources Information Network) adding microbial germplasm fields and strain data into the system: two data base specialists, a curator/nomenclature specialist, and needed software.

$1,000,000/yr

After start-up, costs for adding data and maintaining the GRIN microorganism data would be expected to be relatively stable. Hence, one data specialist and one curator or nomenclature specialist would be needed, along with materials, maintenance of equipment and record keeping. Periodic replacement and upgrades of equipment (hardware, software) are needed.

$500,000/yr

4.2 Central Hub, Strain Repository. Storage facilities at main repository during start up: initial storage of strains, space, equipment, materials for storage.

$500,000/yr

Storage at main repository and periodic shipping of cultures and deposition of new materials.

$250,000/yr

Costs are based on known personnel salaries in the U.S., and minimal needs for a viable germplasm system compared with Canada, the European Union, New Zealand and Japan. The cost are based on a distributed hub and spoke system, with one central hub for retention of all viable cultures, similar to a vault for precious materials, and the distributed Biobanking Centers being the working centers. Costs also are based on using existing facilities primarily, saving considerably in start-up costs, although there may be physical modifications and additions required. Few of these would exceed 3,000 square feet, e.g. a laboratory, two offices and computer equipment area.

The CentraalBureau voor Schimmelcultures (CBS) in the Netherlands, the world's premiere culture collection of fungi, estimates their cost to be 250 Euro per culture for accession, preservation, and storage. The USDA National Center for Agricultural Utilization Research (NCAUR) estimates $40-50 per strain for storage and $20 for distribution. Assuming $50 per strain, it would cost $127,600 to retain cultures in the at risk bacterial collection above. Assuming that is an average size at risk collection we estimate $130,000 per collection.

Transport, collating, computer entry, facility space for long-term storage under liquid nitrogen or -80 refrigeration, and curation of 10 at risk collections at one or more USDA facilities: $130,000 per collection.

Total $1,300,000 to “rescue” 10 at risk collections.

4.3 Specialized Biobank Centers. Database specialist, curator, equipment, hardware, software. 10 centers x $500,000/center/yr = $5 million minimum and 20 centers at $10 million minimum.

Costs would include materials, distribution, tracking, biosafety, quality assurance programs, and some personnel costs. Overhead might be an additional cost. Cost recovery could contribute to a general reduction of the budget, but will add to personnel costs for inventory, record keeping and other duties. Costs would be based on the size of the collection, difficulty of maintaining the strains, difficulty of authentication (identity, pathogenicity or other attributes) and degree of research or collaboration with external scientists on development of methods and organism characterization. Maximum costs:

$25,000,000

4.4 Independent Technical Advisory Committee: An independent technical advisory committee of scientists within and external to the USDA, including the public and private sector, is needed for public support, policies on acquisitions and disbursements, and accountability. Annual or biennial meetings are recommended. Cost for logistical meeting support: 10 members x $2,500 = $25,000/yr.

4.5 Cost Recovery Considerations: The specialized Biobank Centers send out material that is alive or derived from living organisms. Materials are not lent, as in zoological museums or plant herbaria. While the USDA maintains a plant germplasm collection at no cost to the recipient, shipping of living material is relatively costly and time-consuming. This is especially the case for anyone working with Select Agents.

Professional culture collections world-wide, supported primarily by national government funding, charge most requestors a fee for sending a viable specimen. Fees may be waived for some researchers including collaborators in research projects involving the curator or scientific staff or if the material is for educational purposes. A fee helps control the overall budget and deters frivolous requests, both large and small. Additional services might be done for a fee, such as authentication of the strain, strain typing, DNA extraction, safe deposit of strains for special purposes, including industrial strains. With plant pathogens, pathogenicity tests are critical to determine usefulness; some culture collections provide this service, although it may be of a limited scale.

If cost recovery is not considered, then it is estimated by a Canadian report on culture collections, (*National Centres for Secure Biological Resources*, March 31, 2007) that an increase of 20-30% in the governmental base budget should be provided.

Appendix 1. List of potential specialized plant-associated microbe centers: Below are some suggestions for existing specialized centers and the estimated number of strains currently held. Undoubtedly more will be added and criteria must be developed to determine which collections should serve as repositories.

Virus and Viroids
*USDA-ARS Cereal Virus Collection, Fargo, ND: one million stocks
USDA-ARS Potato Virus Y Collection, Ithaca, NY, Aberdeen, ID: 3000 stocks
USDA-ARS, MPPL, Phytoplasma collection, Beltsville, MD

Bacteria
*USDA-ARS, NRRL, Peoria, IL – bacteria including actinomycetes: 19,000 stocks
USDA-ARS, USDA-ARS National Rhizobium Germplasm Collection, Beltsville, MD: 5,000 stocks
*International Collection of Phytopathogenic Bacteria (ICPB), Ft. Detrick, MD
Nine additional USDA-ARS collections with more than 5,000 bacterial strains

Fungi and Fungal-like Organisms
Fungal Genetics Stock Center, Kansas City, MO – Neurospora, Aspergillus, Magnaporthe: 70,000 stocks
*USDA-ARS, NRRL, Peoria, IL – yeast, *Aspergillus, Fusarium, Penicillium:* 60,000 stocks
Fusarium Research Center, Pennsylvania State University, University Park, PA: 17,000 stocks
USDA-ARS Cereal Disease Lab, St. Paul, Minnesota – *Puccinia* and other cereal pathogens: 30,000 samples
USDA-FS, Madison, WI – forest pathogens: 12,000 stocks
USDA-ARS, SMML, Beltsville, MD – plant pathogens and biocontrol agents including *Trichoderma*: 11,000 stocks
*World Phytophthora Genetic Resource Collection (WPC) UC-Riverside, CA: 6500 stocks
USDA-ARS Collection of Entomopathogenic Fungal Cultures (ARSEF), Ithaca, NY: 2400 stocks
*International Culture Collection of (Vesicular) Arbuscular Mycorrhizal Fungi (INVAM), West Virginia University - endomycorrhizal fungi: 1,000 stocks
Phaff Yeast Culture Collection, UC Davis, CA – 800 species of yeasts: 6,000 stocks

Nematodes
*USDA-ARS, Nematology Laboratory, Beltsville, MD – plant-associated nematodes: 12 stocks
Florida *Meloidogyne* Collections – root knot nematodes: 10 stocks

*Denotes Collections at Risk:

Appendix 2: COSTS for NPMGS

Tiered approach

Year One:
Modify and begin populating GRIN database - $1,000,000
Begin specialized center selection and prioritization process $12,500
Initiate Steering Committee process and selection. 0
TOTAL: $1,012,500

Year Two:
Begin *acquisition, vetting and* storage of at risk collections (2) $260,000
GRIN – personnel, etc. (ongoing) $500,000
Add one specialized center as pilot process $500,000
Steering Committee appointment and meeting 25,000
TOTAL: $1,285,000

Year Three:
Continue adding at risk collections and begin backup of center collections (5, 7 total) $650,000
GRIN – central hub (ongoing) $500,000
Add two additional specialized centers (3 total) $1,500,000
Steering Committee activities (meeting) 25,000
TOTAL: $2,675,000

Year Four:
GRIN – personnel $500,000
GRIN – storage $250,000
Add three additional centers (6 total) $3,000,000
Additional at risk collections (3, 10 total) $390,000
Steering Committee activities (meeting) 25,000
TOTAL: $4,165,000

Year Five:
GRIN – personnel $500,000
GRIN – storage $250,000
Add four additional centers (10 total) $5,000,000
Steering Committee (Conference calls) 1,200
TOTAL: $5,751,200

$14,888,700 Total for the first five years
Costs years 5 through 10 $5,750,000/yr x 5 = $28,750,000

Total: first ten years $43,638,700

ANNEX 1

Phytopathology News 2/18/09

Scott Gold, Jeff Jones, Kellye Eversole, and Rick Bennett

Second Workshop to Facilitate the Establishment of a National Plant Microbial Germplasm System (NPMGS)

Under the auspices of the APS-Public Policy Board and the APS Ad-Hoc Committee on Culture Collections and with support and participation from USDA-ARS and USDA-APHIS a Second Workshop to Facilitate the Establishment of a **National Plant Microbial Germplasm System (NPMGS)** was held January 27-28, in Arlington, VA.

Some thirty-five participants, representing APS-PPB, universities, industry, the Fungal Genetics Stock Center, USDA-ARS, USDA-APHIS, the NSF, the Smithsonian Institution, and several international culture collection systems met to define a plan for a NPMGS. Additionally, we were very pleased to have two Congressional staff members from the House of Representatives Committee on Science's Subcommittee on Investigations & Oversight attend and present their perspectives at the meeting. Representative Brad Miller, the chair of the subcommittee is very interested in BioBanking issues, and will reintroduce legislation this year to protect endangered collection resources under Federal stewardship and those supported by Federal funds maintained by non-Federal institutes. The establishment of the NPMGS is intended to provide efficient preservation of, access to, and retrievable documentation for important plant associated microbial resources. The current lack of such a system is widely recognized as limiting to research efforts and potentially dangerous for our capacity to quickly respond to new disease challenges.

A draft NPMGS plan document was generated from a preliminary planning workshop held in 2007. Prior to the second workshop, working subcommittees from amongst the participants addressed challenges in the draft plan related to strategies for the future NPMGS including 1) Management structure; 2) Cultures and collections; 3) System structure and operations; and, 4 Budget needs. Each subcommittee produced brief documents addressing these continuing issues that were shared with all participants prior to the meeting and were presented by the subcommittee co-chairs to initiate the meeting.

The meeting included several presentations by international culture collection and database experts. These presentations, highlighting existing culture collections and related efforts in Europe and Canada, informed participants about the state of the art in database tools for culture collections as well as strategies and funding issues related to the various systems. Importantly, the potential for interconnectivity of a future NPMGS and international collections were discussed at some length. Additionally, a presentation on the current status of the U.S. Interagency Working Group on Scientific Collections was delivered by a representative from the Smithsonian Institution and potential interactions with that group discussed.

Several breakout and roundtable discussions allowed participants to hone ideas generated by the pre-meeting subcommittee reports and the various presentations. These were wide-ranging discussions and they clearly indicated that many detailed issues will need to be resolved by the management of the NPMGS. The idea of taking incremental steps toward building the NPMGS was discussed.

A proposed structure of the NPGMS was more clearly defined through the second workshop. The current envisioned system includes the following core elements: 1) A centralized hub with backup collections located at the USDA-ARS facility at Fort Collins, CO, building on the current structural and expertise resources available through the National Plant Germplasm System (NPGS) and the Germplasm Resource Information Network (GRIN). Investments in additional equipment and personnel would be required to make this possible, but the existing framework of the NPGS is viewed as a valuable way to leverage new investments in the NPMGS. The personnel at the Fort Collins facility are envisioned to be permanent USDA employees.

2) A system of "BioBanks", based primarily on existing collections (government and university, etc.), each with taxon specific expertise.

3) A state of the art, federated database system (likely associated with the GRIN system managed through the NPGS) linking biobank and hub information with connectivity to other domestic and international collection databases to dramatically enhance the usefulness of the collections. More than a catalog, this is envisioned as serving as an integration point for information on collection specifics, genomics, molecular genetics, taxonomy, population genetics, and ecology. Sophisticated user- friendly informatics tools will be incorporated to generate a knowledge rich platform for idea generation.

4) A management structure involving a Steering Committee of federal employees and a Scientific Advisory Board made up of stakeholders, including APS representatives.

The meeting participants agreed to continue as a working group and communication has continued via email. A draft executive summary has been generated and will be provided to the APS PPB for distribution to policy-makers during their mid-year meetings in Washington, D.C. Finally, a white paper more fully describing the NPMGS plan by incorporating ideas generated at the 2nd workshop is underway. The executive summary and the white paper will be posted under Public Policy Initiatives in the Media/Outreach section of APS.net.

Persons with an interest in this issue or who have thoughts about the gaps in the current system or needed components of a NPMGS, are encouraged to contact the Ad Hoc committee chairs at or the PPB JBJones@ufl.edu or jacqueline.fletcher@okstate.edu, respectively.