Plant Pathogen Culture Collections: It Takes a Village to Preserve These Resources Vital to the Advancement of Agricultural Security and Plant Pathology


First to fifth authors: Department of Plant Pathology, The Pennsylvania State University, University Park 16802; sixth author: Department of Geography, The Pennsylvania State University, University Park 16802; seventh author: U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS), Microbial Genomics Research Unit, Peoria, IL 61604; eighth author: USDA-ARS, Foreign Disease-Weed Science Research Unit, Ft. Detrick, MD 21702; ninth author: Pennsylvania Department of Agriculture, Harrisburg, PA 17110; tenth author: Mountain Hort. Crops Research & Extension Center, North Carolina State University, Fletcher 28732; eleventh and twelfth authors: School of Agricultural Biotechnology, Seoul National University, Seoul, Korea; thirteenth author: Horticultural Crops Research Laboratory, USDA-ARS and Department of Botany & Plant Pathology, Oregon State University, Corvallis 97330; fourteenth author: USDA-ARS, 1636 E. Alisal St., CA 93906; fifteenth author: Department of Plant Pathology, University of California, Riverside 92521; and sixteenth and seventeenth authors: Center for Computational Genomics, The Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park 16802.

Accepted for publication 17 April 2006.

ABSTRACT


Plant pathogen culture collections are essential resources in our fight against plant disease and for connecting discoveries of the present with established knowledge of the past. However, available infrastructure in support of culture collections is in serious need of improvement, and we continually face the risk of losing many of these collections. As novel and reemerging pathogen threats increase, our access to rapidly assess these resources is essential for rapid assessment of potential risk and to help track the change and movement of pathogens. Realizing the potential of pathogen genomics as a foundation for developing effective disease control also hinges on how effectively we use the sequenced isolate as a reference to understand the genetic and phenotypic diversity within a pathogen species. In this letter, we propose a number of measures for improving pathogen culture collections.

Additional keywords: database, disease diagnosis, epidemiology, informatics, pathogen taxonomy and systematics, population genetics.

Corresponding author: S. Kang; E-mail address: sxk55@psu.edu

DOI: 10.1094/PHYTO-96-0920
© 2006 The American Phytopathological Society

PHYTOPATHOLOGY
destroy it. Can the plant pathology community afford such losses? The intent of this letter is to foster a community-wide discussion on why and how we should preserve and catalogue plant pathogen cultures.

**WHY IS IT NECESSARY TO BUILD AND PRESERVE CULTURE COLLECTIONS?**

Collections of pathogen isolates form the primary link that connects past, present, and future research endeavors. Such a link provides an essential foundation for basic research, forensics, and risk assessment of newly isolated pathogens. Unavailability or loss of isolates previously characterized often makes it difficult for other scientists to reproduce and expand research based on past studies (17). Because science builds on existing knowledge, the lack of establishing proper links between what has been done and what will be done is a poor scientific practice and frequently forces us to “reinvent the wheel.” Preserving relevant pathogen isolates from past disease epidemics is equivalent to archiving key documents needed for understanding an important historical event. As “those who do not remember the past are condemned to repeat it” (58), our failure to learn from past and present disease epidemics will result in the likelihood of repeatedly suffering from the same epidemic scenario.

**Pathogen diversity in ecosystems remains vastly under-explored.** Application of DNA sequence-based identification techniques (markers such as ribosomal RNA [rRNA]-encoding genes) to study microbial communities in various environments has revealed how little we understand about the diversity of microbes in nature; as much as 99% of bacteria from many environments cannot be isolated using standard culturing techniques (3). Indeed, some bacterial divisions are known to exist only from the sequencing of the 16S rRNA genes from environmental samples (54). 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.

How many species of plant pathogens exist in nature? At the moment, we can only make educated guesses, but the actual number must be very high. For fungi alone, approximately 10,000 species are known as being pathogenic to plants (1,21). Considering the vast magnitude of unexplored fungal diversity (35–37), the actual number of pathogenic fungal species is likely to be much higher than 10,000. Therefore, systematically cataloguing what has already been isolated and studied as references is essential for future exploration.

**Reference cultures with phenotypic and genotypic identification tags can facilitate the identification and control of newly emerging pathogens.** Pathogen taxa are traditionally delineated based on morphological and phenotypic characters. In recent years, DNA-based phylogenetic approaches have been utilized to complement and augment classical systematics because closely related species may be morphologically indistinguishable. Traits within a species can also vary depending on cultural conditions. These problems often make accurate identification of plant pathogens a challenge, even to specialists. With the advent of molecular systematics (26,38,61), it has also become increasingly apparent that plant pathogens once classified as a single species based on morphology are more diverse than previously thought (4,16,28,47,52,53,62). This suggests that molecular identification may be the most effective means for distinguishing pathogen species and populations. As individual reference cultures are permanently linked with genotypic data, utilizing this data for new species descriptions, or accurately identifying isolates of existing species, should help minimize confusion. Genotypic data from archived reference cultures also supports diagnostics and regulatory measures. Diagnosis based on genetic markers unique to a target pathogen is a highly effective tool for detection, mainly due to increased sensitivity and selectivity. In particular, this approach will greatly assist the study of newly isolated pathogens by researchers who have limited experience in pathogen identification, or by regulatory agency scientists who must quickly assess the threat of a new pathogen for rapid deployment of containment and/or eradication measures.

The emergence of new pathogens through hybridization (8–10) and dispersal further underscores the importance of archiving pathogen culture collections that represent the known genotypic and phenotypic diversity and geographical distribution. Because pathogens frequently migrate from one region to another through various means, including agricultural trade, human travel, and weather-related events, pathogen isolates from diverse geographic locations and genotypic data associated with them are crucial for developing effective regulatory and disease management strategies. One example is the ornamental industry, where plants are propagated in various countries and then shipped to the United States; pests and pathogens may accompany this plant material. Thus, the ability to rapidly trace pathogens to its geographic origin based on genetic similarity, as was done for the case of potato late blight (29,31), may help identify a problem at its source. Although it has been several years since *Phytophthora ramorum* was first identified in both the United States and Europe (55,63), its origin remains unknown. It took approximately 5 years to identify and characterize the pathogen (56). If comprehensive data on the genetic diversity of *Phytophthora* species existed on a global scale prior to the outbreak of this disease, subsequent detective work might have been much easier. Considering that *P. ramorum* is unlikely to be the last emerging pathogen, we need to start curating plant pathogen collections and archiving data on the diversity, distribution, and dynamics of pathogens in ecosystems.

DNA sequence-based identification is not limited to microbes. The use of a short DNA sequence as “DNA barcodes” has recently been proposed for the identification of animal and plant species (6,39,43). This proposal led to the formation of an international initiative termed the Consortium for the Barcode of Life. An analogous global initiative to generate and archive genotypic data for plant pathogens is highly desirable. Of course, this does not imply that molecular markers should replace classical phenotypic markers, as genotypic data without biological context has limited value. Both approaches should be considered complementary tools for diagnostics and evolutionary studies of plant pathogens.

Pathogen genomics should be supported by a collection of isolates representing phenotypic and genotypic diversity within sequenced species. Knowing the complete genetic blueprint of pathogens can quickly advance our understanding of pathogen biology via the use of various functional genomic tools (13,25,50,60,64). Genome sequencing and subsequent functional analyses usually concentrate on a single isolate from a given species. However, because crop loss is typically caused by populations of strains that vary in traits such as virulence, host range, compatibility type (e.g., vegetative compatibility and mating), chemical resistance, and/or toxin production, knowing the complete genetic makeup of a single isolate is insufficient for fully understanding the pathogenetic potential within a species. At the genome level, significant variation within a species appears to be the norm. For example, certain bacterial species contain a significant number of genes (up to 25 to 35% of the total genes) that are specific to individual strains (“dispensable genome”), in addition to genes present in all isolates (“core genome”) within the pan genome (core genome plus dispensable genome) of the species (49). Certain pathogenicity genes of fungal pathogens such as *Nectria haematococca* (32) and *Alternaria alternata* (34) are present on a dispensable chromosome and, therefore, are absent in certain isolates of the same species. Thus, realizing the full potential of pathogen genomics as a foundation for developing effective
measures of disease control hinges on how effectively we use genome sequence data to gain a comprehensive understanding of the genetic and phenotypic diversity within pathogen species across temporal and geospatial scales. Therefore, parallel efforts to archive and characterize collections of isolates representing diverse phenotypic, spatial, and temporal diversity for the sequenced species are essential.

**HOW SHOULD PATHOGEN CULTURES BE ARCHIVED AND IMPROVED?**

In this section, we propose a number of considerations in archiving pathogen culture collections and associated information. Properly archiving cultures and data associated with them should be considered as important as collecting and analyzing new accessions. The lack of proper mechanisms and protocols to support archival and allow access of available cultures will significantly diminish their value and utility. This can lead to confusion about taxonomic identity and duplication of research. Therefore, building both physical and informatics infrastructures to support the cataloguing, access, and use of culture collections is essential.

**Preservation of pathogen cultures is not sufficient.** Whenever possible, each archived isolate should be linked to its geographic origin, host/substrate of origin, date collected, phenotypes (such as cultural and developmental morphology, host range, disease symptoms, recombination strategy/type, chemical resistance, and/or toxin production), genotypic data, and resulting publications in a digitized format that can be easily accessed and searched. Genotyping is routinely performed in many laboratories and, in comparison to generating phenotypic data, is not time or labor intensive. Similarly, collection of data on geographic and host origin can be done easily and inexpensively, especially with the availability of affordable global positional system (GPS) devices. Unfortunately, many pathogen cultures archived to date lack detailed information on their geographic origin, typically ending at the country or state/province level. This hinders efforts to gain insight into the temporal and spatial dynamics of the species of interest. Cataloguing the geographic origin of isolation of cultures will also permit access and use of the digital archives of geographic and environmental data, such as (i) biophysical coverage (http://nationalmap.usgs.gov), (ii) remote sensing data resources (http://terraref.wr.usgs.gov/resource.html), and (iii) climate and weather data (http://www.noaa.gov/), to better understand factors affecting disease and pathogen dynamics.

If we catalogue the types of data described above for archived strains, the value and utility of the culture collection would improve in a number of ways: (i) supporting the identification of a newly isolated pathogen via the use of genetic markers, (ii) assessing the potential risk the new isolate poses (e.g., virulence, toxigenicity, pesticide resistance, exotic nature, etc.) based on its genetic similarity to those isolates that have been characterized previously at the phenotypic level, (iii) building a global map of phenotypic and genotypic diversity and distribution within individual pathogen species, and (iv) facilitating the development of sensitive and selective molecular diagnostic tools. Such data would also facilitate the efficient use of storage space at culture collection facilities. In a perfect world, we would save all the isolates that are being collected, but in reality, this is unrealistic. It is, therefore, critical to establish a mechanism whereby a decision on which isolates to be saved is evaluated scientifically. The ability to compare the phenotypic and genotypic data from new isolates with those from existing isolates should help the selection process.

**Existing research and survey initiatives should be used to enhance the content and utility of pathogen culture collections.** To address agricultural security concerns, the U.S. Department of Agriculture (USDA) has launched a nationwide disease monitoring system, termed the National Plant Diagnostic Network (NPDN), by linking and upgrading disease diagnostic laboratories in land-grant institutions. The NPDN can potentially serve as a conduit for collecting important pathogen isolates and associated data at the national level. Another opportunity is increased research interest and survey activities in response to specific pathogens. In response to the threat from *P. ramorum*, multiple federal and state agencies in the United States, as well as other agencies in Canada and the European Union, have been surveying forests and nurseries. One of the by-products of such surveys is the isolation of many *Phytophthora* isolates from a broad range of hosts and geographical and ecological areas (15,40). Not surprisingly, some of these isolates turned out to be new species (33,41,42). Cultures and data from these surveys offer a golden opportunity for mapping the spatial and temporal diversity and dynamics of *Phytophthora* spp. on a global scale. To support the cataloguing and use of genotypic and phenotypic data from the genus *Phytophthora*, the *Phytophthora* Database project (www.phytophthoradb.org) was launched in 2005.

The National Science Foundation (NSF)-sponsored “Assembling the Fungal Tree of Life (AFTOL)” project (45) is a global effort to address long standing phylogenetic questions within the kingdom Fungi by using DNA sequences of multiple phylogenetically informative loci and morphological traits in approximately 1,500 species representing all major fungal clades. Because plant pathogens are not monophyletic, a comprehensive phylogenetic framework from this project will provide an excellent foundation for studying the evolution of plant pathogenic fungi and will also assist in classifying and identifying new pathogen species. The USDA-ARS Systematic Botany and Mycology Laboratory (SBML), which maintains the U.S. National Fungus Collections, also offers a comprehensive resource for fungal pathogens. The data available at the SBML include descriptions, nomenclature and illustrations of plant-associated fungi, reports of fungi on plants based on the literature, and a database of literature references for the identification of plant-associated fungi, which provides a historical overview on the known geographic distribution and hosts of a specific pathogen species (21). The examples described above are not the only resources we can use to enhance the content and utility of pathogen culture collections. If properly curated, much of the data published in plant pathology-related journals can also be a source.

**A cohesive cyber-infrastructure should support the management and use of culture collections.** Information science provides computational tools, protocols, and standards that can help scientists integrate databases and share methods and research findings. Federated databases and ontology-based database integration provide tools for coordinated data management and archiving; portal technology and web services provide customizable access points to methods and data; and grid technology coupled with high-speed data connections provide distributed but high-powered computational resources (12,20,30,57,59). A coordinated approach to describing, archiving, and disseminating research data via the use of such computational resources can yield significant gains in productivity and the quality of research (59).

Given the need for systematically cataloguing a diverse array of pathogen species/populations and associated data into a format that can be easily accessed and updated, developing a cohesive computational environment (cyber-infrastructure), should be an integral part of enhancing culture collections. In addition to the database that will house the genotypic and phenotypic data, we envision the following data mining and visualization tools to ensure full utilization of the stored data: (i) web interfaces for data submission, search, and visualization; (ii) genotype-based search tools such as BLAST (2) and multi-locus sequence typing (11) to allow for identification of an unknown isolate to the closest species or population by querying the database (19,27);
leverage each other’s research. It will also increase the pace at culture collections but also for enabling distributed teams of epidemiology. Therefore, creating such a cyber-infrastructure is tools will only grow as we accumulate more data. In the long run, recombination.

conducive to the emergence of new variants of pathogens via distribution patterns of specific pathogens and (ii) high risk areas logical factors, such as (i) the likely cause(s) of spatially variant gain a more holistic picture of pathogen dynamics and epidemiological factors, such as the geographic origins of the isolates to vast digital archives of environmental, geospatial and agronomic data at the given locations (23). In combination, these tools will allow users of the culture collection to visualize chosen pathogen species and populations in evolutionary, geospatial and environmental contexts and help them gain a more holistic picture of pathogen dynamics and epidemiological factors, such as the geographic origins of the isolates to vast digital archives of environmental, geospatial and agronomic data at the given locations (23). In combination, these tools will allow users of the culture collection to visualize chosen pathogen species and populations in evolutionary, geospatial and environmental contexts and help them gain a more holistic picture of pathogen dynamics and epidemiological factors, such as (i) the likely cause(s) of spatially variant distribution patterns of specific pathogens and (ii) high risk areas vulnerable to disease outbreak by the selected pathogens or conducive to the emergence of new variants of pathogens via recombination.

The importance and utility of data mining and visualization tools will only grow as we accumulate more data. In the long run, these tools will serve as a bridge between pathogen ecology and epidemiology. Therefore, creating such a cyber-infrastructure is essential not only for the long-term growth and viability of the culture collections but also for enabling distributed teams of researchers and regulators to collaborate more closely so as to leverage each other’s research. It will also increase the pace at which new knowledge and warning is transmitted. Producing and preserving genomic DNA from archived isolates should be an integral part of any culture collection. One major benefit of preserving DNA from archived cultures is that the strains can be characterized later using different genetic markers and tools as new research questions emerge. Rapidly accumulating genome sequence data will provide a wealth of new markers (such as genome-wide single nucleotide polymorphisms and simple sequence repeats) for comprehensive population genetic studies. Progress in pathogen genomics will also lead to the identification of many genes controlling the traits of agricultural importance. Reanalyzing previously characterized isolates using genetic markers that are tightly linked to such genes will increase our ability to predict the phenotypic traits of a newly isolated pathogen through genotyping and would be greatly facilitated by the availability of DNA samples. Pathogen genomics has also stimulated the development of powerful genotyping tools such as micro-/macroarray-based methods for identifying multiple pathogens and detecting a large number of nucleotide polymorphisms (7,22,44,48). Such tools are likely to increase the interest and need for reanalyzing previously characterized isolates. Increasing opportunities for employing comprehensive genotyping analyses to study pathogen biology underscore the importance of cataloguing and rigorously verifying phenotypic data from the archived isolates (such as pathogenicity, chemical resistance, and toxin production). Genotypic data that are not backed by an accurate biological context have limited value and may even create confusion.

Due to the increased stringency of regulations regarding possession, shipment, and use of cultures in research, there is a clear need in the plant pathology community for the availability of DNA extracts from archived reference isolates. For those interested in population genetics and phylogenetics, acquiring DNA of appropriate reference cultures, rather than the cultures themselves, should significantly speed up the progress of their studies by allowing them to bypass the permit acquisition, culturing, and DNA preparation steps. Because of the availability of whole genome amplification methods such as multiple displacement amplification (18,24), the amount of DNA that needs to be preserved and shipped can be quite small.

CONCLUSIONS

In response to the urgency of protecting existing pathogen culture collections, we suggest a number of measures. However, these measures mainly address the issue of how to improve their value and utility and are not sufficient; other critical issues remain to be addressed by the plant pathology community as a whole to ensure the sustainability of culture collections. For example, how do we secure and manage long-term financial resources for maintaining and improving culture collections? Although it may not be difficult to convince federal and state policy makers about the importance of culture collections for agricultural security, given other worthy competing needs under a severe budget constraint, creating a new resource requires more than good justification.

The plant pathology community, in consultation with curators and hosting institutions of culture collections and the federal agencies supporting and/or regulating them, needs to develop a comprehensive and uniform management plan. Given the strict regulations on the use and movement of pathogen cultures, this management plan should address the issue of how to balance research needs and security concerns. Other concerns, such as the ownership of cultures and the timely deposition of cultures used in published research into appropriate culture collection facilities, should also be addressed. Although many culture collections have been established by dedicated individuals, managing them as personal collections is not a desirable strategy for their long-term sustainability, and may discourage the deposition of cultures by other scientists.

Given the vast diversity of pathogens in relation to the dwindling number of trained field pathologists and systematic biologists, a long-term investment in related training with support from academic institutions and government agencies cannot be overemphasized. 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. Without adequate human capital, the future of culture collections and our preparedness for agricultural security will be jeopardized.

“Know the enemy and know yourself; in a hundred battles you will never be in peril. When you are ignorant of the enemy but know yourself, your chances of winning or losing are equal. If ignorant both of the enemy and of yourself, you are certain in every battle to be in peril” (from the Art of War written by the Chinese philosopher Sun Tzu and translated by Samuel B. Griffith). Although many challenges lie ahead in reaching the point that plant disease problems are proactively managed based on a solid understanding of pathogen diversity and biology, judicious use of existing resources and emerging technologies, and better cooperation and coordination within the global plant pathology community should facilitate the establishment of a virtually linked global pathogen culture and data archive. Once established, this resource will serve as a central hub for facilitating the sharing, comparison, and analysis of data, and will establish a baseline for monitoring the emergence of new/foreign pathogens in temporal, geospatial, and environmental contexts.
Due to its vast scope and need for community participation, this effort will only succeed if it is founded on community support and cooperation.

ACKNOWLEDGMENTS

We thank the USDA-NRI Plant Biosecurity program (2005-35605-15393) for its support for the *Phytophthora* Database project. This letter is dedicated to the late L. J. Franchl (1949–2005), a respected colleague, mentor, leader, and scientist.

LITERATURE CITED


