

The Promise and Pitfalls of Sequence-Based Identification of Plant-Pathogenic Fungi and Oomycetes

Seogchan Kang, Michele A. Mansfield, Bongsoo Park, David M. Geiser, Kelly L. Ivors, Michael D. Coffey, Niklaus J. Grünwald, Frank N. Martin, C. André Lévesque, and Jaime E. Blair

First to fourth authors: Department of Plant Pathology, The Pennsylvania State University, University Park 16802; fifth author: Mountain Hort. Crops Research & Extension Center, North Carolina State University, Mills River 28759; sixth author: Department of Plant Pathology and Microbiology, University of California, Riverside 92521; seventh author: Horticultural Crops Research Laboratory, USDA-ARS and Department of Botany & Plant Pathology, Oregon State University, Corvallis 97330; eighth author: USDA-ARS, Salinas, CA 93906; ninth author: Agriculture and Agri-Food Canada, Ottawa K1A 0C6, Canada; and tenth author: Department of Biology, Franklin & Marshall College, Lancaster, PA 17604.

Accepted for publication 10 April 2010.

ABSTRACT

Kang, S., Mansfield, M. A., Park, B., Geiser, D. M., Ivors, K. L., Coffey, M. D., Grünwald, N. J., Martin, F. N., Lévesque, C. A., and Blair, J. E. 2010. The promise and pitfalls of sequence-based identification of plant-pathogenic fungi and oomycetes. *Phytopathology* 100:732-737.

Sequences of selected marker loci have been widely used for the identification of specific pathogens and the development of sequence-based diagnostic methods. Although such approaches offer several advantages over traditional culture-based methods for pathogen diagnosis and identification, they have their own pitfalls. These include erroneous and incomplete data in reference databases, poor or oversimplified inter-

pretation of search results, and problems associated with defining species boundaries. In this letter, we outline the potential benefits and drawbacks of using sequence data for identification and taxonomic deductions of plant-pathogenic fungi and oomycetes, using *Phytophthora* as a primary example. We also discuss potential remedies for these pitfalls and address why coordinated community efforts are essential to make such remedies more efficient and robust.

Additional keywords: culture collection, database, diagnosis, DNA barcode, phylogeny, taxonomy.

Within the past decade plant pathologists have become increasingly dependent on publicly accessible sequence databases such as GenBank and EMBL to identify organisms based on DNA sequence similarity. This application has been particularly important in instances where users may not possess the expertise or reference material necessary to make accurate identifications using morphological and cultural characteristics. The latter approach, while absolutely vital in many instances, may not be practical in cases where time is of the essence in the prevention, control, or eradication of serious plant disease outbreaks. From a human resources standpoint, those interpreting morphological and cultural characteristics often need years of training and experience in order to make an accurate identification. From a systematics standpoint, DNA sequences are highly objective because they are an intrinsic characteristic of an organism, not dependent on environmental or cultural conditions, and are highly reproducible. Naturally, multiple efforts to develop DNA markers that can support phylogenetic analysis and/or rapid identification have emerged, covering everything from particular groups of organisms (e.g., genera, ecological groups such as mycorrhizal fungi) to entire kingdoms (3,8,9,33). Besides supporting pathogen identification and providing robust phylogenetic frameworks for understanding the evolution of pathogens, resulting data have aided the development and validation of molecular diagnostic tools (32,44,50).

The Assembling the Fungal Tree of Life project represents the global fungal community effort to generate and curate phylogenetic and phenotypic data from a wide range of species representing the major branches of the Kingdom Fungi (9,33). The Consortium for the Barcode of Life (<http://www.barcoding.si.edu/>) is an international organization that aims to tag individual known species in all kingdoms of life with unique DNA barcode sequences (4,23,46). This approach has enabled rapid species identification and discovery and has also aided in creating biodiversity inventories (26,43). It is important to point out that tree of life initiatives and DNA barcoding are often complementary. The markers for tree of life studies are sequenced for a few representative species across an entire kingdom or phylum and should align well, whereas barcode sequences are typically generated for all species in a genus and do not necessarily have to align well throughout the genus, families or orders. Barcoding represents the twigs and leaves of the tree, whereas tree of life initiatives provide information about the branches (13). Examples of the species identification platforms with a barcode-like approach include the FUSARIUM-ID (<http://www.fusariumdb.org/>), a *Fusarium* identification database (20); the *Phytophthora* Database (<http://www.phytophthoradb.org/>) (39) and *Phytophthora*-ID (<http://phytophthora-id.org/>), databases supporting the identification and molecular systematics of *Phytophthora*; UNITE (<http://unite.ut.ee/index.php>), an internal transcribed spacer (ITS) database supporting the identification of ectomycorrhizal fungi (28); TrichOKey (<http://www.isth.info/tools/molkey/index.php>), a database supporting the identification of *Hypocrea* and *Trichoderma* species (15); and BOLD (<http://www.boldsystems.org>), the main barcode of life data system (41) which now contains an extensive database of ITS and cytochrome oxidase 1 (COI) for oomycetes.

First, second, and third authors contributed equally to this work.

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

doi:10.1094/PHYTO-100-8-0732

© 2010 The American Phytopathological Society

The best examples of such databases harbor expert-curated DNA sequences attached to publicly available, vouchered cultures, providing users with data and taxonomic identifications that are trustworthy at best and scientifically disprovable at worst. With recent advances in molecular methodologies, many of the processes required for sequence generation have become faster, cheaper, and less labor intensive. Accordingly, DNA-based methods for systematics and species identification will continue to improve and expand in plant pathology. However, these methods are not the proverbial silver bullets of modern taxonomy or pathogen diagnosis, as they, like all other methods, have their own limitations. Therefore, unless certain precautions are taken in generating and interpreting data, the resulting answers may be misleading.

POTENTIAL PITFALLS IN SEQUENCE-BASED PATHOGEN IDENTIFICATION

Erroneous and incomplete data in public databases. The BLAST (basic local alignment search tool) program (1) infers relatedness between sequences based on the statistical significance of matches and has been widely used to identify organisms. Because this approach relies on sequences from previously characterized organisms for comparison, the quality of the resulting answers is only as good as the existing reference data set. As such, the lack of sequence data from identical or closely related taxa can lower the probability of an accurate identification. However, because of the rapid increase in sequence data from a wide range of genera, especially of the ITS region of ribosomal RNA encoding genes, the use of this marker as a query generally leads to sequences of closely related species. The ITS region has been proposed as the primary barcode for fungi and oomycetes (45,46). There were 2,425 ITS sequences derived from *Phytophthora* isolates in GenBank in April 2008; by January 2009, these had increased by 10.1% (Table 1). Although the number of ITS sequences representing individual *Phytophthora* species varied widely, with more than 50% of the sequences corresponding to just 10 species, the majority of described and recently discovered species appear to be represented in GenBank.

Public databases like GenBank contain a small yet significant fraction of erroneous and/or incomplete data (6,7,22,31,34,49). Although GenBank manually curates the submitted data to ensure the accuracy of locus annotation, taxonomic accuracy is validated only at higher taxonomic levels. Therefore, if the submitted data are generated from a misidentified isolate, it subsequently becomes associated with incorrect species information and is difficult to rectify because only the depositor of sequences is allowed to make changes. In cases where the taxon information has changed or is in flux (e.g., *P. parasitica* = *P. nicotianae*; *P. arecae* = *P. palmivora*, *P. sinensis* = *P. melonis*), the old information usually remains in GenBank. In addition, a large number of sequences in GenBank are annotated as "*Phytophthora* sp." (Table 1), suggesting that the depositors were unclear about species identity or considered them to be putative new species at the time of data deposition. To determine whether the sequences labeled "*Phytophthora* sp." actually represent potentially novel species that await formal description, we compared them with other sequences available in GenBank and the *Phytophthora* Database (39) via BLAST. Approximately 33% of the sequences labeled "*Phytophthora* sp." were identical to the ITS sequences from described species (data not shown), illustrating that many of the "*Phytophthora* sp." accessions are incompletely annotated.

To assess the type and degree of potential errors associated with *Phytophthora* in GenBank, we downloaded all *Phytophthora* ITS sequences in April 2008 and aligned them based on their species annotation. We limited our analysis to the accessions that cover the whole ITS region (i.e., those in the 701 to 900 bp range) and excluded the sequences labeled as "*Phytophthora* sp." Conse-

quently, the total number of ITS sequences analyzed was 1,940. Sequences that did not align well with those from other strains within the same species were then subjected to BLAST searches against GenBank and the *Phytophthora* Database. This analysis revealed that 25 accessions were clearly misidentified, as their sequences showed much higher identity to sequences from distantly related species than to the annotated species designation. Some of these errors were likely caused by morphological similarity between certain species. For instance, *P. melonis* was traditionally considered a synonym of *P. drechsleri* (24), but this proposal has not been upheld with subsequent molecular analyses (3,12). Not surprisingly, six accessions labeled as *P. drechsleri* (AF228093, AF228095, AY251650, AY251651, AY251652, and DQ501284) had much higher sequence homology with *P. melonis* sequences. Because we focused on sequences that clearly differed from others within data sets of individual species and omitted those that had higher sequence similarity to closely related species, the number of misidentified *Phytophthora* accessions is likely much higher than 25.

A significant number of *Phytophthora* ITS accessions in GenBank are only partial sequences of the ITS region (e.g., 262 and 330 accessions being shorter than 700 bp in the 2008 and 2009 data sets, respectively) with some being shorter than 200 bp (62 and 70 accessions in the 2008 and 2009 data sets, respectively). While some researchers may choose to analyze shorter portions of marker loci depending on their research objectives, deposition and use of partial sequences can potentially contribute to confusion in pathogen identification, especially in the taxon groups that have not been comprehensively and accurately characterized (see below). Human mistakes during editing and annotation have been documented in public sequence databases (6,22). In addition, some sequence errors can also result from problems during the amplification or sequencing steps of data generation (11).

Difficulties in interpreting BLAST search results. The top score match may not necessarily represent the same species or even a close relative, especially when the diversity of the target taxon is poorly represented in the database. BLAST scores and e-values can be influenced by factors other than the evolutionary relationships between the query and search results; one such factor is the length of the alignment made between the match and the query, labeled as the query coverage. A perfect match with one of the numerous partial sequences mentioned above will have a much lower score than a longer alignment with the incorrect species that has several base pair differences. It is very common to obtain part of the small subunit (SSU) and large subunit (LSU)

TABLE 1. The number and distribution of internal transcribed spacer (ITS) sequences from *Phytophthora* species in GenBank^a

Species	April 2008	January 2009
<i>P. capsici</i>	197	198
<i>P. cactorum</i>	164	177
<i>P. cinnamomi</i>	155	162
<i>P. cryptogea</i>	149	152
<i>P. infestans</i>	144	146
<i>P. citricola</i>	134	153
<i>P. nicotianae</i>	111	169
<i>P. megasperma</i>	92	95
<i>P. citrophthora</i>	83	83
<i>P. ramorum</i>	75	77
<i>Phytophthora</i> sp.	152	213
Other species	945 (90)	1,040 (91)
Total number	2,425	2,670

^a All ITS sequences available in GenBank were downloaded twice (April 2008 and January 2009) and grouped according to their species annotation. *Phytophthora* sp. includes ITS sequences without species identity. Other species corresponds to all ITS sequences from species other than the top 10 and those labeled *Phytophthora* sp. The number of species in the other species group is shown in parentheses.

of rRNA genes from the sequencing of the ITS region. For some environmental studies of oomycetes, long reads from SSU or LSU are deliberately obtained in addition to ITS in order to generate at least a genus-level match. Sending a BLAST query with both ITS and LSU will return accessions that have both LSU and ITS at the top, i.e., the sequences that have the longest matching sequence coverage to the query will have the highest score.

For many organisms, relying on information from a single gene region, most commonly ITS, narrows the possible identity to the genus and occasionally, but not always, to the species level (8,12,18,29). However, the use of sequence identity at a single locus to conclude identity at the species level can be problematic as there is no reasonable threshold of sequence identity that justifies species boundaries. Also, a single gene may not have comparable rates of evolution from one species to another. Speciation within the genus *Phytophthora*, like most organisms, has resulted from a variety of factors, including but not limited to geospatial separation, host selection, and/or mating isolation (5). Therefore, defining species boundaries based on the nature and rate of sequence change in a single locus is problematic, as different genes can tell different stories. Multilocus-based species recognition provides a means to dissect species boundaries among closely related taxa, particularly when a single locus fails on its own to distinguish them (19,48).

In *Phytophthora*, comparison of ITS sequences may reveal little difference between closely related but distinct species, while at the same time exhibiting a high degree of variation within certain species. Our in-depth phylogenetic analyses of *P. cinnamomi* and *P. capsici* (M. A. Mansfield, D. M. Geiser, Gugino, F. N. Martin, J. E. Blair, S. Kang, and M. D. Coffey, unpublished data) and the data generated to construct the *Phytophthora* Database (3,39) illustrate the complexity of this issue. We have sequenced the ITS regions for nearly 100 isolates each of *P. cinnamomi* and *P. capsici*. In *P. cinnamomi*, all isolates were identical at this locus, despite the wide host and geographical range represented. However, in *P. capsici* there was considerable variation with the degree of sequence difference between some *P. capsici* isolates being even greater than that between several closely related species (e.g., *P. cactorum* versus *P. hedraiaandra* and *P. infestans* versus *P. phaseoli*). Additional multigene phylogenetic analyses suggest that traditionally defined *P. capsici* is a complex consisting of several cryptic species (M. A. Mansfield, D. M. Geiser, Gugino, F. N. Martin, J. E. Blair, S. Kang, and M. D. Coffey, unpublished data). For some species, the ITS region is invariable as exemplified with *P. infestans* relatives (17). A MEGABLAST search (using *P. ipomoeae* query AY770742.1) resulted in equally good hits to *P. infestans* and *P. mirabilis*.

As already mentioned, it is not uncommon to encounter isolates that exhibit some level of intraspecific variation, although this is generally on the order of one to several base pairs. Among closely related species this can be the same level of variation seen on an interspecific level thereby resulting in multiple species having the same BLAST score. One approach to reduce the impact of this problem is to evaluate a wide range of isolates to identify those portions of the ITS region that are highly conserved within a species yet different from the others and conduct the sequence-based identification using only these conserved regions; this has been done for *Trichoderma* spp. (15). The barcode of life data system (BOLD) alleviates this problem by using exactly the same partial gene region for all reference sequences. These sequences align perfectly and a neighbor joining analysis is immediately available following the BLAST search to compare the query with the matched data.

Erroneous data (e.g., the 25 *Phytophthora* accessions described above along with human/technical errors) and incomplete data (i.e., sequences annotated as "*Phytophthora* sp." and very short sequences) contribute to confusing BLAST results, but these pitfalls are relatively easy to recognize. However, erroneous infor-

mation gets amplified when users simply rely on the top BLAST match as the final determinant of an organism's identification and subsequently deposit their data with the same erroneous annotation. The DNA barcode approach alleviates some of these problems by giving the "DNA barcode reference" label only if electropherograms and primer sequences are deposited, if the specimens or strains are properly vouchered and accessible, and if the sequence is of proper length. This provides an opportunity for others to verify the quality of the sequences and perform additional work if there are any taxonomic issues. However, currently the GenBank barcode label is only available for COI sequences from animals, and the best barcode region for fungi or oomycetes remains to be determined and agreed upon (45,46).

Hybrid species. *P. alni* is thought to be a hybrid species consisting of three subspecies (25). Similarly, *P. andina* may be a hybrid between *P. infestans* and another unknown parent, possibly *P. mirabilis* (21). Sequences from a hybrid *Phytophthora* species are often only provided as a single sequence read using IUB (International Union of Biochemistry) codes for heterozygous sites, or worse, a consensus sequence excluding heterozygous sites, thus removing the hybrid signal. To identify hybrid taxa, cloning followed by sequencing are thus necessary. Current databases do not easily accommodate separate sequences cloned from a hybrid taxon, and mitochondrial sequence databases will be unable to detect hybrids due to the uniparental inheritance of this organelle.

Discord between previously described species based on morphological traits and other species concepts. This pitfall is illustrated by the species complex that was observed among morphologically identified isolates of *P. capsici*. The genus *Phytophthora*, like many microorganisms, has a limited number of phenotypic characters, many of which can be affected by cultural conditions. To further complicate matters, the interpretation of these characters can be subjective, as illustrated by changes over time in *Phytophthora* taxonomy (16,18,30,42,47). The advent of DNA-based techniques has allowed us to examine the existence of cryptic species, which are unlikely to be easily distinguished based on morphological and cultural traits. We are in a period in pathogen taxonomy where data based on different species concepts (e.g., biological, morphological, and phylogenetic) coexist and conflict sometimes. This can lead to problems with correct identification when relying strictly on BLAST results (particularly for other species complexes such as *P. citricola*, *P. citrophthora*, *P. cryptogea*, *P. drechsleri*, and *P. megasperma* that have yet to be fully resolved). Besides *Phytophthora*, the taxonomy of the genus *Fusarium*, which has long been a model of taxonomic instability, also demonstrates the issue. Certain traditionally defined species, such as *F. graminearum* and *F. solani*, include many distinct phylogenetic species (35,38). We suspect that as more data become available for lesser-studied pathogenic groups, similar patterns will likely emerge. Unfortunately, discord between classical and molecular taxonomy has not been well studied in most pathogen groups. Even for those cases that have been examined, the resulting data and potential conflicts have not been fully resolved and are not widely recognized by the plant pathology community.

HOW TO NAVIGATE THROUGH THE PITFALLS?

Coordinated community action in the generation and archiving of high quality data. All BLAST results should be interpreted with care in determining species identity, and in some cases additional data and analysis may be required. Once sequence database users clearly understand the nature and source of potential pitfalls and limitations, most of the problems in sequence-based identification can be avoided or resolved. However, in the long run, the best remedy is generating high quality data (e.g., accurate sequence barcodes, new species descriptions

founded on a robust phylogenetic framework) based on vouchered cultures, and archiving both the data and cultures in a format that can be easily searched and retrieved as community references. Addressing existing conflicts in species resolution comprehensively is also critical. To make these efforts successful, coordinated community action is essential; a piecemeal or isolated case approach can potentially create confusion instead of the order that taxonomy should provide. Potential problems could be managed more rigorously in a tailored organism-specific community database (8,39,49), resulting in a much clearer taxonomic framework of the target pathogen groups, and also contribute to accurately documenting the earth's biodiversity through the organisms' genotypes as envisioned by the Barcode of Life initiative (43). Coordination and integration of community projects to analyze and document the genotypic and phenotypic diversity of plant pathogens and their nonpathogenic relatives will help (i) to build solid and comprehensive reference phylogenetic frameworks for future research and education, (ii) to utilize most effectively limited resources, and (iii) to document pathogen diversity worldwide. Plant diseases are an inherently global problem because pathogens frequently migrate from one region to another through various means. In order to respond to pathogen threats based on accurate diagnostics, collaboration to document and describe pathogen species should be global in scale.

One example that illustrates the value of proactively and collaboratively building such community resources is how the recent global outbreak of fusarial keratitis (10), corneal infections caused by *Fusarium*, associated with the use of a particular brand of contact lens solution, was handled. Because of the previous phylogenetic studies on *Fusarium* isolates associated with human infections (37,51), investigators were able to quickly identify multiple species of *Fusarium* associated with eye infections, and determined that they were most likely introduced from the patient environment and not by product contamination. Without the proactive analysis of *Fusarium* collections, these investigations would have taken many months to complete rather than a few days, resulting in greater human and financial harm. Because of such proactive studies based on well-curated culture collections at the Fusarium Research Center, the CBS-KNAW Fungal Biodiversity Center, the Agricultural Research Service (NRRL) Culture Collection, and the Canadian Collection of Fungal Cultures (CCFC/DAOM), *Fusarium* went from being one of the most taxonomically unstable fungal genera to a model of fungal molecular phylogenetic systematics.

For *Phytophthora*, the availability of genus-wide, multi-gene data, including ITS sequences, within the *Phytophthora* Database website (3,39) has facilitated the description of numerous new species and the characterization of species complexes. The compilation of comprehensive data sets from traditional and molecular taxonomic studies into this platform has also helped in the recognition of errors and inconsistencies in the existing data, which in turn has directed the elimination of these problems and the development and testing of new phylogenetic hypotheses. Recently discovered novel *Phytophthora* species of global significance, such as *P. ramorum* and *P. kernoviae*, are unlikely to be the last *Phytophthora* threats to agricultural and forest ecosystems, highlighting the importance of providing such means for data storage, sharing, and utilization to support monitoring the diversity, distribution and dynamics of *Phytophthora* worldwide.

A logical extension of community efforts to provide molecular and bioinformatics tools for species identification is to build community infrastructures that support identification of individuals within species. Because plant diseases are caused by populations of strains that may vary in multiple traits such as virulence, host range, fungicide resistance, and toxin production, we often need more than species identity to develop disease management strategies. Comprehensive data on the genetic and phenotypic structure within species will help us understand genotype-phenotype

relationships and the mechanisms by which pathogen populations change in response to host/environmental selection pressures. Some communities have already begun collaboratively collecting detailed data on spatial and temporal structures of populations within species to build a global picture; a notable example is EuroBlight (<http://www.euroblight.net/EuroBlight.asp>). The goal of this European Union-funded project is to examine variation in European *P. infestans* populations by standardizing and collating the wealth of genotypic and phenotypic data collected by co-operators in >30 participating countries via a single comprehensive database. This database now includes data from South and Central American countries and becomes a powerful reference resource for monitoring and understanding global *P. infestans* populations. More initiatives of this magnitude are needed for pathogens of global economic importance.

The rapid accumulation of sequenced microbial genomes, driven by increasing sequencing capacity and decreasing costs, will allow sequence-based species identification and population genotyping to become easier and more robust. One method for genotypically analyzing populations is multilocus sequence typing. This method provides an unbiased means by which isolates can be identified based on molecular signatures, even in cases where the taxonomy remains unsettled. For example, the *F. solani* species complex consists of dozens of phylogenetically recognizable species, but the majority of these are undescribed or taxonomically vague, and may remain so for a long time. In the meantime, a simple nomenclatural system which identifies isolates as coming from a particular phylogenetic species (designated by a numeral) and a multilocus haplotype within a species (designated by an alphabetic system) has been established in the *F. solani* species complex for easy identification (36) and allows matches at the intraspecific level.

Preservation of key reference cultures. This effort should accompany improved community coordination in data generation and sharing. The construction of the *Phytophthora* Database and accompanying phylogenetic studies rapidly progressed because most of the key cultures representing the known diversity of *Phytophthora* species and their genomic DNA were readily available in the World *Phytophthora* Genetic Resource Collection. Similarly, the genus *Fusarium* has become by far the most well characterized pathogen group due to the existence of well-preserved culture collections. Pathogen germplasm collections are irreplaceable and invaluable resources because they support the scientific discovery process as the primary link between discoveries of the past, present and future (14,27,45). Loss of previously studied isolates often prevents other scientists from confirming and building upon past research and forces us to spend years reestablishing the necessary groundwork to understand economically important pathogens.

CONCLUSIONS

Here we have reviewed the current status of sequence-based species identification and some of the current pitfalls in using this approach, as well as suggesting a few preventative steps to avoid these pitfalls. Although we have mainly focused on *Phytophthora* to highlight potential pitfalls associated with pathogen identification and taxonomic placement, these issues are not unique to this genus. Previous reports have pointed out a number of self-propagating errors and problems that large public databases are prone to, including sequences associated with misidentified organisms, the inability to easily change or update data, and the lack of association or inability to link sequence data to voucher material for verification (6,22,34,49).

Since it is unrealistic, even undesirable, to expect GenBank to curate the taxonomic information for individual accessions, improving the quality of data in GenBank falls primarily on users who deposit and analyze the data. The use of GenBank's

“barcode” keyword only for those accessions that meet a minimum set of criteria is a good step in the right direction but is currently available only to the animal community. Although the idea of manual data curation by certain GenBank users was proposed to improve the quality of data (2,40), acceptance of this proposal by GenBank has not been forthcoming as it would require a significant alteration of its role (40). Of course, with some species being well defined at the morphological and molecular levels, all searches of publicly accessible sequence databases to identify pathogens do not always require the most rigorous quality checks and/or in-depth phylogenetic analysis, and GenBank represents a tremendous global resource as long as users are aware of the potential pitfalls.

The real challenge we are facing as a community is coordinating and supporting researchers who work on individual pathogen groups in building robust and comprehensive data, material, and data sharing infrastructures. As noted earlier, high quality data can eliminate most of the current problems associated with sequence-based identification, however, such data will only quickly become available if we work closely together. Given severe budget constraints in many countries, creating new resources to support such work in a global scale will be challenging. However, considering the potential benefits from such investment and the urgent need for supporting better pathogen diagnosis and monitoring to protect global agricultural and environmental systems, waiting for a better financial environment is not a sound approach; this calls for immediate global brainstorming to develop a road map.

ACKNOWLEDGMENTS

We acknowledge the USDA-AFRI Plant Biosecurity Program for its support for the *Phytophthora* Database project (2005-35605-15393 and 2008-55605-18773). The work of C. A. Lévesque is partly supported by the Canadian Barcode of Life Network.

LITERATURE CITED

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
2. Bidartondo, M. I. 2008. Preserving accuracy in GenBank. *Science* 319:1616.
3. Blair, J. E., Coffey, M. D., Park, S.-Y., Geiser, D. M., and Kang, S. 2007. A multi-locus phylogeny for *Phytophthora* utilizing markers derived from complete pathogen genomes. *Fungal Genet. Biol.* 45:266-277.
4. Blaxter, M. 2003. Counting angels with DNA. *Nature* 421:122-124.
5. Brasier, C. M., and Hansen, E. M. 1992. Evolutionary biology of *Phytophthora*. *Annu. Rev. Phytopathol.* 30:173-200.
6. Bridge, P. D., Roberts, P. J., Spooner, B. M., and Panchal, G. 2003. On the unreal ability of published DNA sequences. *New Phytol.* 160:43-48.
7. Bridge, P. D., Spooner, B. M., and Roberts, P. J. 2004. Reliability and use of published sequence data. *New Phytol.* 161:15-17.
8. Bruns, T. D., Szaro, T. M., Gardes, M., Cullins, K. W., Pan, J. J., Taylor, D. L., Horton, T. R., Kretzer, A., Garbelotto, M., and Li, Y. 1998. A sequence database for the identification of ectomycorrhizal basidiomycetes by phylogenetic analysis. *Mol. Ecol.* 7:257-272.
9. Celio, G. J., Padamsee, M., Dentinger, B. T. M., Bauer, R., and McLaughlin, D. J. 2006. Assembling the fungal tree of life: Constructing the structural and biochemical database. *Mycologia* 98:850-859.
10. Chang, D. C., Grant, G. B., O'Donnell, K., Wannemuehler, K. A., Noble-Wang, J., Rao, C. Y., Jacobson, L. M., Crowell, C. S., Sneed, R. S., Lewis, F. M. T., Schaffzin, J. K., Kainer, M. A., Genese, C. A., Alfonso, E. C., Jones, D. B., Srinivasan, A., Fridkin, S. K., and Park, B. J. 2006. Multistate outbreak of *Fusarium keratitis* associated with use of a contact lens solution. *JAMA* 296:953-963.
11. Clark, A. G., and Whittam, T. S. 1992. Sequencing errors and molecular evolutionary analysis. *Mol. Biol. Evol.* 9:744-752.
12. Cooke, D. E. L., Drenth, A., Duncan, J. M., Wagels, G., and Brasier, C. M. 2000. A molecular phylogeny of *Phytophthora* and related oomycetes. *Fungal Genet. Biol.* 30:17-32.
13. Crandall, K. A., and Buhay, J. E. 2004. Genomic databases and the tree of life. *Science* 306:1144-1145.
14. Crous, P. W., and Cothier, E. J. 2003. The reality of a virtual laboratory: Reference to personal herbaria and culture collections is bad science.

- Inoculum 54:2-3.
15. Druzhinina, I. S., Kopchinskiya, A. G., Komoja, M., Bissett, J., Szakacs, G., and Kubiceka, C. P. 2005. An oligonucleotide barcode for species identification in *Trichoderma* and *Hypocrea*. *Fungal Genet. Biol.* 42:813-828.
16. Erwin, D. C., and Ribeiro, O. K. 1996. *Phytophthora* Diseases Worldwide. The American Phytopathological Society, St. Paul, MN.
17. Flier, W. G., Grünwald, N. J., Kroon, L. P. N. M., van den Bosch, T. B. M., Garay-Serrano, E., Lozoya-Saldaña, H., Bonants, P. J. M., and Turkensteen, L. J. 2002. *Phytophthora ipomoeae* sp. nov., a new species causing leaf blight on *Ipomoea longipedunculata* in the Toluca Valley of central Mexico. *Mycol. Res.* 106:848-856.
18. Gallegly, M. E., and Hong, C. 2008. *Phytophthora*: Identifying Species by Morphology and DNA. The American Phytopathological Society, St. Paul, MN.
19. Geiser, D. M. 2004. Practical fungal molecular taxonomy. Pages 1-12 in: *Advances in Fungal Biotechnology for Industry*. J. Tkacz and L. Lange, eds. Kluwer Academic Publishers, Dordrecht, the Netherlands.
20. Geiser, D. M., Jiménez-Gasco, M., Kang, S., Makalowska, I., Veeraraghavan, N., Ward, T. J., Zhang, N., Kuldau, G. A., and O'Donnell, K. 2004. FUSARIUM-ID v.1.0: A DNA sequence database for identifying *Fusarium*. *Eur. J. Plant Pathol.* 110:473-479.
21. Gomez-Alpizar, L., Hu, C.-H., Oliva, R., Forbes, G., and Ristaino, J. B. 2008. Phylogenetic relationships of *Phytophthora andina*, a new species from the highlands of Ecuador that is closely related to the Irish potato famine pathogen *Phytophthora infestans*. *Mycologia* 100:590-602.
22. Harris, D. J. 2003. Can you bank on GenBank? *Trends Ecol. Evol.* 18:317-319.
23. Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* 270:313-321.
24. Ho, H. H. 1986. *Phytophthora melonis* and *P. sinensis* synonymous with *P. drechsleri*. *Mycologia* 78:907-912.
25. Ios, R., Andrieux, A., Marcaisa, B. T., and Frey, P. 2006. Genetic characterization of the natural hybrid species *Phytophthora alni* as inferred from nuclear and mitochondrial DNA analyses. *Fungal Genet. Biol.* 43:511-529.
26. Janzen, D. H., Hallwachs, W., Blandin, P., Burns, J. M., Cadiou, J.-M., Chacon, I., Dapkey, T., Deans, A. R., Epstein, M. E., Espinoza, B., Franclemont, J. G., Haber, W. A., Hajibabaei, M., Hall, J. P. W., Hebert, P. D. N., Gauld, I. A. N. D., Harvey, D. J., Hausmann, A., Kitching, I. A. N. J., Lafontaine, D. O. N., Jean-Francois, Claude, L., Jacqueline, L., Miller, Y., Miller, J. S., Miller, L. E. E., Miller, S. E., Montero, J., Munroe, E., Green, S. R. A. B., Ratnasingham, S., Rawlins, J. E., Robbins, R. K., Rodriguez, J. J., Rougerie, R., Sharkey, M. J., Smith, M. A., Solis, M. A., Sullivan, J. B., Thiaucourt, P., Wahl, D. B., Weller, S. J., Whitfield, J. B., Wood, K. R. W. D. M., Woodley, N. E., and Wilson, J. J. 2009. Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Mol. Ecol. Res.* 9(s1):1-26.
27. Kang, S., Blair, J. E., Geiser, D. M., Khang, C., Park, S., Gahegan, M., O'Donnell, K., Luster, D. G., Ivors, K. I., Kim, S. H., Lee, Y., Lee, Y., Grünwald, N. J., Martin, F. M., Coffey, M. D., Veeraraghavan, N., and Makalowska, I. 2006. Plant pathogen culture collections: It takes a village to preserve these resources vital to the advancement of agricultural security and plant pathology. *Phytopathology* 96:920-925.
28. Koljalg, U., Larsson, K.-H., Abarenkov, K., Nilsson, R. H., Alexander, I. J., Eberhardt, U., Erland, S., Hoiland, K., Kjöller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A. F. S., Tedersoo, L., Vralstad, T., and Ursing, B. M. 2005. UNITE: A database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytol.* 166:1063-1068.
29. Lee, S. B., White, T. J., and Taylor, J. W. 1993. Detection of *Phytophthora* species by oligonucleotide hybridization to amplified ribosomal DNA spacers. *Phytopathology* 83:177-181.
30. Leonian, L. H. 1925. Physiological studies on the genus *Phytophthora*. *West Virginia Agric. Exp. Stn. Sci. Paper* 11:444-498.
31. Lévesque, C. A., and de Cock, A. W. A. M. 2004. Molecular phylogeny and taxonomy of the genus *Pythium*. *Mycol. Res.* 108:1363-1383.
32. Lievens, B., Brouwer, M., Vanachter, A. C. R. C., Lévesque, C. A., Cammue, B. P. A., and Thomma, B. P. H. J. 2005. Quantitative assessment of phytopathogenic fungi in various substrates using a microarray. *Environ. Microbiol.* 7:1698-1710.
33. Lutzoni, F., Kauff, F., Cox, C. J., McLaughlin, D., Celio, G., Dentinger, B., Padamsee, M., Hibbett, D., James, T. Y., Baloch, E., Grube, M., Reeb, V., Hofstetter, V., Schoch, C., Arnold, A. E., Miadlikowska, J., Spatafora, J., Johnson, D., Hambleton, S., Crockett, M., Shoemaker, R., Sung, G.-H., Lücking, R., Lumbsch, T., O'Donnell, K., Binder, M., Diederich, P., Ertz, D., Gueidan, C., Hansen, K., Harris, R. C., Hosaka, K., Lim, Y.-W., Matheny, B., Nishida, H., Pfister, D., Rogers, J., Rossman, A., Schmitt, I., Sipman, H., Stone, J., Sugiyama, J., Yahr, R., and Vilgalys, R. 2004.

- Assembling the fungal tree of life: Progress, classification, and evolution of subcellular traits. *Am. J. Bot.* 91:1446-1480.
34. Nilsson, R. H., Ryberg, M., Kristiansson, E., Abarenkov, K., Larsson, K.-H., and Koljalg, U. 2006. Taxonomic Reliability of DNA Sequences in Public Sequence Databases: A Fungal Perspective. *PLoS One* 1:e59.
 35. O'Donnell, K. 2000. Molecular phylogeny of the *Fusarium solani*/*Nectria haematococca* species complex. *Mycologia* 92:919-938.
 36. O'Donnell, K., Sutton, D. A., Fothergill, A., McCarthy, D., Rinaldi, M. G., Brandt, M. E., Zhang, N., and Geiser, D. M. 2008. Molecular phylogenetic diversity, multilocus haplotype nomenclature, and in vitro antifungal resistance within the *Fusarium solani* species complex. *J. Clin. Microbiol.* 46:2477-2490.
 37. O'Donnell, K., Sutton, D. A., Rinaldi, M. G., Magnon, K. C., Cox, P. A., Revankar, S. G., Sanche, S., Geiser, D. M., Juba, J. H., van Burik, J.-A. H., Padhye, A., Anaissie, E. J., Francesconi, A., Walsh, T. J., and Robinson, J. S. 2004. Genetic diversity of human pathogenic members of the *Fusarium oxysporum* complex inferred from multilocus DNA sequence data and amplified fragment length polymorphism analyses: Evidence for the recent dispersion of a geographically widespread clonal lineage and nosocomial origin. *J. Clin. Microbiol.* 42:5109-5120.
 38. O'Donnell, K., Ward, T. J., Geiser, D. M., Kistler, H. C., and Aoki, T. 2004. Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genet. Biol.* 41:600-623.
 39. Park, J., Park, B., Veeraraghavan, N., Blair, J. E., Geiser, D. M., Isard, S., Mansfield, M. A., Nikolaeva, E., Park, S.-Y., Russo, J., Kim, S. H., Greene, M., Ivors, K. L., Balci, Y., Peiman, M., Erwin, D. C., Coffey, M. D., Jung, K., Lee, Y.-H., Rossmann, A., Farr, D., Cline, E., Grünwald, N. J., Luster, D. G., Schrandt, J., Martin, F., Ribeiro, O. K., Makalowska, I., and Kang, S. 2008. *Phytophthora* database: A cyber infrastructure supporting the identification and monitoring of *Phytophthora*. *Plant Dis.* 92:966-972.
 40. Pennisi, E. 2008. DNA DATA: Proposal to 'Wikify' GenBank meets stiff resistance. *Science* 319:1598-1599.
 41. Ratnasingham, S., and Hebert, P. D. N. 2007. BOLD: The barcode of life data system: Barcoding. *Mol. Ecol. Notes* 7:355-364.
 42. Rosenbaum, J. 1917. Studies of the genus *Phytophthora*. *Proc. Natl. Acad. Sci. USA* 3:159-163.
 43. Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K., and Lane, R. 2005. Towards writing the encyclopedia of life: An introduction to DNA barcoding. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360:1805-1811.
 44. Schaad, N. W., Frederick, R. D., Shaw, J., Schneider, W. L., Hickson, R., Petrillo, M. D., and Luster, G. D. 2003. Advances in molecular-based diagnostics in meeting crop biosecurity and phytosanitary issues. *Ann. Rev. Phytopathol.* 41:305-324.
 45. Seifert, K. A. 2008. Integrating DNA barcoding into the mycological sciences. *Inoculum* 59:2-6.
 46. Seifert, K. A. 2009. Progress towards DNA barcoding of fungi. *Mol. Ecol. Res.* 9(s1):83-89.
 47. Stamps, D. J., Waterhouse, G. M., Newhook, F. J., and Hall, G. S. 1990. Revised tabular key to the species *Phytophthora*. *Mycol. Pap.* 162:1-28.
 48. Taylor, J. W., Jacobson, D. J., Kroken, S., Kasuga, T., Geiser, D. M., Hibbett, D. S., and Fisher, M. C. 2000. Phylogenetic species recognition and species concepts in fungi. *Fungal Genet. Biol.* 31:21-32.
 49. Vilgalys, R. 2003. Taxonomic misidentification in public databases. *New Phytol.* 160:4-5.
 50. Vincelli, P., and Tisserat, N. 2008. Nucleic acid-based pathogen detection in applied plant pathology. *Plant Dis.* 92:660-669.
 51. Zhang, N., O'Donnell, K., Sutton, D. A., Nalim, F. A., Summerbell, R. C., Padhye, A. A., and Geiser, D. M. 2006. Members of the *Fusarium solani* species complex that cause infections in both humans and plants are common in the environment. *J. Clin. Microbiol.* 44:2186-2190.