



HAL
open science

PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes

Nalvo Almeida, Shuangchun Yan, Rongman Cai, Christopher R. Clarke, Cindy E. Morris, Norman W. Schaad, Erin L. Schuenzel, George H. Lacy, Xiaonan Sun, Jeffrey B. Jones, et al.

► **To cite this version:**

Nalvo Almeida, Shuangchun Yan, Rongman Cai, Christopher R. Clarke, Cindy E. Morris, et al.. PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. *Phytopathology*, 2010, 100 (3), pp.208-215. 10.1094/PHYTO-100-3-0208 . hal-02668768

HAL Id: hal-02668768

<https://hal.inrae.fr/hal-02668768>

Submitted on 31 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PAMDB, A Multilocus Sequence Typing and Analysis Database and Website for Plant-Associated Microbes

Nalvo F. Almeida, Shuangchun Yan, Rongman Cai, Christopher R. Clarke, Cindy E. Morris, Norman W. Schaad, Erin L. Schuenzel, George H. Lacy, Xiaolan Sun, Jeffrey B. Jones, Jose A. Castillo, Carolee T. Bull, Scotland Leman, David S. Guttman, João C. Setubal, and Boris A. Vinatzer

First, second, third, fourth, eighth, and sixteenth authors: Department of Plant Pathology, Physiology, and Weed Science, Virginia Tech, Blacksburg 24061; first and fifteenth authors: Virginia Bioinformatics Institute, Virginia Tech; first author: Department of Computing and Statistics, Federal University of Mato Grosso do Sul, Brazil; fifth author: Unité de Pathologie Végétale UR407, INRA, F-84140 Montfavet, France; sixth and seventh authors: United States Department of Agriculture–Agricultural Research Service (USDA-ARS), Foreign Disease–Weed Science Research, Ft. Detrick, MD 21702-5023; ninth author: Department of Agriculture and Consumer Services, Division of Plant Industry, Gainesville, FL 32608; tenth author: University of Florida, Plant Pathology Department, P.O. Box 110680, Gainesville 32611; eleventh author: PROINPA Foundation, P.O. Box 4285, Cochabamba, Bolivia; twelfth author: USDA-ARS, Salinas, CA 93905; thirteenth author: Department of Statistics, Virginia Tech; fourteenth author: Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario, M5S3B2 Canada; and fifteenth author: Department of Computer Science, Virginia Tech., Blacksburg 24061.

Accepted for publication 3 November 2009.

ABSTRACT

Almeida, N. F., Yan, S., Cai, R., Clarke, C. R., Morris, C. E., Schaad, N. W., Schuenzel, E. L., Lacy, G. H., Sun, X., Jones, J. B., Castillo, J. A., Bull, C. T., Leman, S., Guttman, D. S., Setubal, J. C., and Vinatzer, B. A. 2010. PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. *Phytopathology* 100:208-215.

Although there are adequate DNA sequence differences among plant-associated and plant-pathogenic bacteria to facilitate molecular approaches for their identification, identification at a taxonomic level that is predictive of their phenotype is a challenge. The problem is the absence of a taxonomy that describes genetic variation at a biologically relevant resolution and of a database containing reference strains for comparison. Moreover, molecular evolution, population genetics, ecology, and epidemiology of many plant-pathogenic and plant-associated bacteria are still poorly understood. To address these challenges, a database with web

interface was specifically designed for plant-associated and plant-pathogenic microorganisms. The Plant-Associated Microbes Database (PAMDB) comprises, thus far, data from multilocus sequence typing and analysis (MLST/MLSA) studies of *Acidovorax citrulli*, *Pseudomonas syringae*, *Ralstonia solanacearum*, and *Xanthomonas* spp. Using data deposited in PAMDB, a robust phylogeny of *Xanthomonas axonopodis* and related bacteria has been inferred, and the diversity existing in the *Xanthomonas* genus and in described *Xanthomonas* spp. has been compared with the diversity in *P. syringae* and *R. solanacearum*. Moreover, we show how PAMDB makes it easy to distinguish between different pathogens that cause almost identical diseases. The scalable design of PAMDB will make it easy to add more plant pathogens in the future.

Additional keywords: bacterial plant pathogens, plant disease diagnostics.

The current taxonomy of many plant-pathogenic and plant-associated bacteria presents a serious obstacle to the identification of bacterial plant pathogens and to the discrimination of closely related strains with different host-specific virulence potential (7,42). Many named species of bacterial plant pathogens are heterogeneous and comprise organisms that are genetically and phenotypically diverse and have a variety of hosts. This is due, in part, to changes in nomenclature that were dictated by the International Code of Nomenclature of Prokaryotes (24). For example, many species in the genera *Pseudomonas* or *Xanthomonas* were lumped into *Pseudomonas syringae* or *Xanthomonas campestris*, respectively, in 1980. Their names were not included in the Approved Lists of Bacterial Names (7,36) because their description was either incomplete or no type strain was available to

serve as a reference to which future comparisons could be made. In order to maintain information about these plant pathogens, the former specific epithets became subspecific pathovar designations, indicating that the previous nomenclatures had distinct host ranges or disease reactions (12,46,48). For example, the organism that causes bacterial speck disease of tomato was formerly called *P. tomato* and became known as *P. syringae* pv. *tomato* (12).

Adding to the heterogeneity of strains of a given pathovar are the minimal criteria used in assigning strains to previously existing taxa. In some cases, strains were assigned to pathovars based only on the host from which they were isolated without further host range testing and little phenotypic or genetic characterization (7,46). Therefore, strains given the same pathovar designation may belong to phylogenetically distinct groups and strains given different pathovar designations may be phylogenetically identical.

Several taxonomic proposals have elevated *Xanthomonas* pathovars back to species since 1980 (23,35,41) based on a combination of data, including DNA-DNA hybridization (DDH) and phenotypic characterization (i.e., using polyphasic taxonomy) (42). DDH is an indirect measurement of relatedness that allows the distinction between species based on the “operational species

Corresponding author: B. A. Vinatzer; E-mail address: vinatzer@vt.edu

*The e-Xtra logo stands for “electronic extra” and indicates that the online version contains three figures showing neighbor-joining trees of the genus *Xanthomonas*, *Pseudomonas syringae*, and *Ralstonia solanacearum* and a table showing the *Xanthomonas* strains used in this study.

doi:10.1094/PHYTO-100-3-0208

© 2010 The American Phytopathological Society

concept" (i.e., two bacteria belong to different species if their measured DDH value is <70%, <5% ΔT_m , and biochemical or other phenotype-based tests can distinguish between them) (37,44). However, DDH results depend heavily upon the choice of experimental methods (S_1 nuclease method [21] versus optical method [11]) and experimental parameters that cannot be completely standardized among laboratories or experiments. Therefore, two bacteria that have $\approx 70\%$ DDH may or may not be considered to be the same species based on slight differences in the experimental set up. A significant problem with DDH is that, when new bacteria are analyzed, reference strains must be included in the experiment, thus increasing the number of assays and cost. Moreover, reciprocal hybridization experiments need to be completed because they may give different results. More generally, the description of bacterial species based on the operational species concept may have little biological relevance and reveals very little about genetic diversity within a bacterial species or in regard to life cycle or ecological niche occupied by members of the species.

Within *P. syringae sensu lato*, only two pathovars have been elevated to species since 1980. Gardan and colleagues (15) described nine genomospecies based on DDH but only proposed that two, *P. tremae* and *P. cannabina*, be elevated to named species because they were the only genomospecies with unique phenotypic differences that would allow for discrimination from the other genomospecies. Thus, there was phylogenetic evidence for elevating additional species but no straightforward methods for allocating strains to these species.

Ralstonia solanacearum is a third group of plant-pathogenic bacteria with problematic classification and nomenclature and, thus, difficult identification. First, races were used to describe groups within *R. solanacearum* with different host ranges; then, biovars were introduced to distinguish between groups biochemically; finally, phylotypes were introduced to distinguish between groups based on phylogeny (4). In addition to the lack of correspondence between the different classification methods, no formal nomenclature has been proposed for any of these classifications and, thus, no type strains have been proposed to represent the different races, biovars, and phylotypes.

Multilocus sequence typing (MLST) was first described in 1998 as a bacterial typing approach (27) taking advantage of the qualitative and transportable nature of DNA sequence data (i.e., the result of DNA sequencing of a particular strain is the same regardless of the technology that was used to obtain the sequence and regardless of the laboratory where the experiment was performed) (26). Moreover, if sequences are deposited in a public database, they can be viewed and compared with sequences of other strains from any computer with Internet access. 16S ribosomal (r)RNA sequencing is also qualitative and transportable and has been included in the description of many species. However, the limitation of 16S rRNA sequencing is that the sequences are not diverse enough to distinguish between closely related bacterial species (see, for example, research by Hauben and colleagues in regard to *Xanthomonas* spp.) (18). MLST consists of the sequencing of multiple loci (typically four to eight "housekeeping" genes; that is, genes conferring a basic metabolic function) and the chosen loci are much more diverse in nucleotide composition than 16S rRNA sequences; therefore, MLST usually allows strains to be distinguished below the species level (as defined by the operational species concept). Moreover, sequencing several loci facilitates further investigation of molecular evolution and population genetics parameters; for example, the relative contribution of mutation and recombination to the evolution of a species (29).

In an MLST study, each allele sequence at each locus is assigned an allele number and the combination of alleles at the analyzed loci is referred to as sequence type (ST). For example, strains assigned to ST 1 may have alleles designated with number

1 at locus A through locus F whereas strains with ST 2 may have the same alleles as strains belonging to ST 1 at five of the six loci but may have a new allele, designated (for example) allele 2 at locus B. MLS analysis (MLSA) instead focuses on comparisons of the actual DNA sequences (16). MLST and MLSA studies have been published for several plant-pathogenic bacteria: *P. syringae* (20,28,32,43,45), *Xanthomonas* spp. (1,47), *R. solanacearum* (8), *Xylella* spp. (33), *P. viridiflava* (17), and *Acidovorax citrulli* (14). These studies were mainly performed to elucidate evolution and population genetics of these pathogens. In *P. syringae*, these studies have also allowed clarification of the relationship between agricultural strains and strains from nonagricultural and nonplant substrates, such as snow, rain, and lakes (28). In *Xanthomonas* spp., MLSA results were analyzed in regard to species that had been previously described based on DDH results (47).

For many human pathogens, MLST data have been deposited by users around the world in curated, publicly accessible databases that contain hundreds or even thousands of strains or species (for example, for *Neisseria meningitidis*) (5). In addition to the genetic information, geographic location and time of isolation are also included in these databases for each strain. This makes it possible not only to study evolution and population genetics but also to investigate disease epidemiology; for example, to identify disease outbreaks caused by new variant strains (although this is not always possible because new variant strains may differ from other strains in only a small number of virulence genes not assayed by MLST). The use of DNA sequences or STs deposited electronically in public databases makes these studies independent of nomenclature. Moreover, laboratories in different countries and continents can compare their strains with each other without the constraint of transporting the strains or the limitations imposed by quarantine regulations.

In this article, an MLST/MLSA database and website for plant-associated and plant-pathogenic microbes named the Plant-Associated Microbes Database (PAMDB) is described. PAMDB allows MLST/MLSA data and critical metadata of strains to be deposited and searched. Metadata include location and year of isolation and images of the substrate of origin, whereby we define a substrate as a specific plant, seed, snow, rain, or soil sample from which bacteria were isolated at a specified location on a specified date. We demonstrate how PAMDB can be used to clarify the relationships among phytopathogens with a newly developed *Xanthomonas* MLST scheme. We also present results of a comparison of the diversity in terms of genetic distance between *P. syringae*, *Xanthomonas* spp., and *R. solanacearum* using data deposited in PAMDB. Finally, examples of the use of PAMDB to distinguish between different pathogens causing almost identical diseases are shown.

MATERIALS AND METHODS

Bioinformatics. PAMDB consists of four main components: a relational database built with the open source database software MySQL (<http://www.mysql.org>), a set of scripts written in the programming language Perl (<http://www.perl.org>), a BLAST server (<http://www.ncbi.nlm.nih.gov/Ftp/>), and static HTML pages, all of them working under Apache Server (<http://www.apache.org>). The Perl modules CGI and DBI, both available from CPAN (<http://www.cpan.org>), are used. The user interface is managed by 43 Perl scripts. The scripts generate the HTML forms that the users interact with, manage the session, accept and process requests, preserve website security and users preferences, and dynamically display the results in HTML. The user interface look and feel is primarily defined by standard cascading style sheets (CSS). Google maps (<http://maps.google.com>) are embedded dynamically in PAMDB using Perl scripts and using a Java Application Programming Interface available from the

Google Maps website (<http://code.google.com/apis/maps/>). The whole system is hosted on a Dell server running Ubuntu Linux (<http://www.ubuntu.com/>).

Experimental methods. Degenerate primers were designed for the *Xanthomonas* genes *fusA*, *gapA*, *gltA*, *gyrB*, *lacF*, and *lepA*. Allele sequences for these genes were obtained from the National Center for Biotechnology Information for the sequenced genomes of *X. campestris* American Type Culture Collection (ATCC) 33913 (10), *X. axonopodis* subsp. *citri* strain 306 (10), and *X. oryzae* pv. *oryzae* KACC10331 (25). Sequences were aligned in Seqman (DNA*, Madison, WI). Regions (50 to 100 bp long) with high sequence identity between the three sequenced genomes at an approximate distance of 400 to 600 bp from each other were chosen as locations for forward and reverse primers. *X. axonopodis* subsp. *citri* strain 306 sequences corresponding to these regions were used for primer design in primer3 (<http://frodo.wi.mit.edu/primer3/>) and manually degenerated to make them anneal to the allele sequences of *X. campestris* ATCC 33913 and *X. oryzae* pv. *oryzae* KACC10331. Primer sequences and polymerase chain reaction PCR conditions are listed in Table 1. Genomic DNA was extracted from xanthomonads as previously described (35). PCR and sequencing of PCR products were performed as described previously (44).

Molecular evolution and populations genetics. MEGA 4 (40) was used to estimate genetic distances shown in Table 2. Standard errors were computed applying the bootstrap method with 10,000 replicates and randomly generated seed (also in MEGA4).

The Bayesian tree shown in Figure 1 was generated in MrBayes 3.1.2 using the Markov Chain Monte Carlo (MCMC) method. The evolution model was set to general time reversible (GTR) with gamma-distributed rate variation across sites and a proportion of invariable sites. The program was run for 30,000,000

generations, which was long enough to ensure the standard deviation of split frequencies to be <0.01. Sample frequency was 150. When summarizing the substitution model parameters and trees, 50,000 samples were used for burn-in. The Bayesian tree was rooted with *Stenotrophomonas maltophilia* R551 3 as out-group in TreeView PPC 1.6.6 (<http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>).

Neighbor-joining (NJ) trees were constructed in PAUP 4.0b10 (39) for *R. solanacearum* with 1,000 bootstrap replicates and randomly generated starting seed. The branch-swapping algorithm of tree-bisection-reconnection (TBR) was employed during bootstrap. Through swapping and comparing between trees, the best unrooted NJ trees were selected. NJ trees were constructed in Megalign (DNA*) for *Xanthomonas* spp. and *P. syringae* performing 1,000 bootstrap trials with a starting seed of 111.

RESULTS

PAMDB database structure. PAMDB consists of a relational database and a website. The database is composed of 11 tables that contain metadata of substrates, metadata and sequence data of strains, and user data (Table 3 provides a complete description of each table). The website (<http://genome.ppws.vt.edu/MLST>, also accessible through the URL www.pamdb.org) is composed of pages that are dynamically built based on database content as well as of static HTML pages (for example, the homepage and help documents). Employing dynamically built pages allows adding new organisms, new loci, or entire new MLST schemes to the PAMDB database without creating any new web pages and by making only minimal manual changes to some of the static pages (for example, to update the help documents with new primer information).

TABLE 1. Loci of the *Xanthomonas* multilocus sequence typing and analysis scheme and primers used for polymerase chain reaction and sequencing

Locus name	Fragment length (bp)	Primer sequences	Annealing temperature (°C)	Final primer concentration (µM)
fusA	591	For TCTGGCSCARGARGAYCC	58	0.6
		Rev GCCTCTTCGTARTGGTCRAA	58	0.6
gap-1	444	For GGCAATCAAGGTTGGYATCAACG	58	0.4
		Rev ATCTCCAGGCACTTGTTSGARTAG	58	0.6
gltA	501	For ATCTTGATCAGGTCACGCTCAAC	58	0.2
		Rev AGCATCTTCAGCACGGCTTCGTT	58	0.2
gyrB	411	For AAGTTCGACGACAACAGCTACAA	58	0.2
		Rev GAMAGCACYCGGATCATGCCTTC	58	0.2
lacF	408	For SAGRITCCACCACTTGAAGC	58	0.6
		Rev SAGRITCCACCACTTGAAGC	58	0.2
lepA	390	For AAGCSCAGGTGCTCGACTCCAAC	58	0.2
		Rev CGTTCCTGCACGATTTCCATGTG	58	0.2

TABLE 2. Mean and maximum distance between strains of *Pseudomonas syringae*, *Ralstonia solanacearum*, and *Xanthomonas* spp. and phylogroups therein

Group	No. of STs ^a	Mean distance ^b		Maximum distance ^c
		Jukes-Cantor	Tajima-Nei	
<i>P. syringae</i>	84	0.066 (0.004)	0.067 (0.004)	0.106
Phylogroup I	18	0.016 (0.002)	0.016 (0.002)	0.053
Phylogroup II	33	0.030 (0.002)	0.031 (0.002)	0.06
Phylogroup III	24	0.020 (0.002)	0.020 (0.002)	0.042
<i>R. solanacearum</i>	35	0.028 (0.002)	0.028 (0.002)	0.051
Phylogroup I	14	0.005 (0.001)	0.005 (0.001)	0.011
Phylogroup II	11	0.016 (0.001)	0.016 (0.001)	0.028
Phylogroup IV	6	0.007 (0.001)	0.007 (0.001)	0.013
<i>Xanthomonas</i> spp.	51	0.047 (0.002)	0.048 (0.002)	0.077
Phylogroup I	8	0.006 (0.001)	0.006 (0.001)	0.015
Phylogroup II	32	0.026 (0.002)	0.026 (0.002)	0.042
Phylogroup IIa	8	0.004 (0.001)	0.004 (0.001)	0.006
Phylogroup IIb	6	0.003 (0.001)	0.003 (0.001)	0.007
Phylogroup IIc	5	0.008 (0.001)	0.008 (0.001)	0.011

^a ST = sequence type; number of nonredundant (unique) allele combinations per group.

^b Mean distance and standard errors thereof were calculated in Mega4 (40).

^c Maximum pairwise distance was calculated in Megalign (DNA*, Madison, WI).

Data query and data extraction. To browse and search data in PAMDB, users need to register at the website by providing an e-mail address, name, and affiliation, and choose a password. At login, users enter their username and password and select the organism they want to work with from a pull-down menu. Once logged in, users can browse or search strains and substrates by keyword. Users can view metadata for strains, including map location and images of the substrates from which the strains came from (as long as the submitter of these strains provided that information) (Table 3 provides a complete list of strain and substrate data that can be provided).

Before using PAMDB for strain identification, users will first have to determine the pathogen group the strain probably belongs to by employing phenotypic tests or 16s rRNA sequencing. Once the strain has been identified—for example, as belonging to the genus *Xanthomonas*—users will sequence one or several loci of their unknown strain using *Xanthomonas* spp.-specific primers. Primer sequences are listed in the PAMDB help pages online. Using protocols also available at the website, users can obtain DNA sequences for the chosen loci. Sequences can either be compared individually to the locus sequences already deposited in PAMDB using the BLAST program (3) or users can create a file with all their sequences and BLAST them against the database in “batch mode”. However, each sequence in the file will be used in a separate BLAST search and, therefore, each hit in the database

will be based on similarity with a single sequence in the file. Details on how to perform a BLAST search are also available online. The BLAST result page will list hits with links to strain and substrate metadata and an embedded Google map showing the geographic origin of strains (if locations were provided by submitters).

Because PAMDB has already been populated with many sequences from common *P. syringae*, *R. solanacearum*, and *Xanthomonas* pathogens, users may find that the sequences of their strain are 100% identical to sequences in the database. If these sequences all belong to strains of the same pathogen—for example, *P. syringae* pv. *tomato*—this provides a very strong support that the user’s strain belongs to *P. syringae* pv. *tomato*. However, to confirm that result, at least one more locus sequence should give a 100% hit to the same pathogen. If the sequences of the user’s strain do not match any sequence in the database 100%, the user can download all sequences from PAMDB (see next paragraph), align the sequences of their own strain with the sequences downloaded from PAMDB, and build a phylogenetic tree using third-party software to place their strain relative to all strains in PAMDB. Of course, users can also choose to submit metadata and sequence data of their strain to PAMDB (see next section), which will allow other users to compare their strains to the user’s strain.

Users can also view allele profiles for all strains deposited in PAMDB and download individual allele sequences or con-

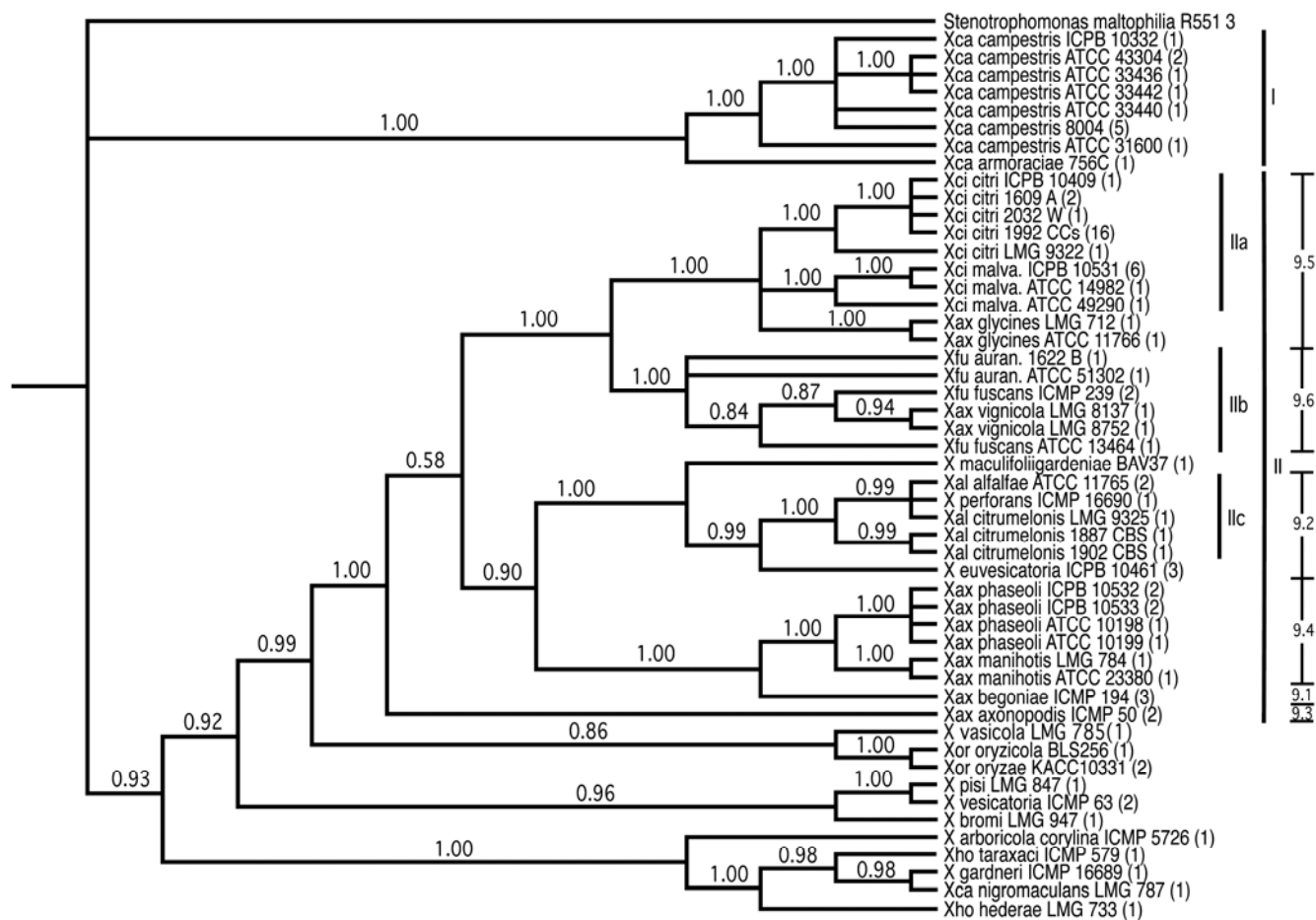


Fig. 1. Phylogenetic tree of the genus *Xanthomonas* based on the concatenated data set of fragments of the genes *fusA*, *gap-1*, *gltA*, *gyrB*, *lacF*, and *lepA*. The tree is a consensus tree constructed with MrBayes (31). *Stenotrophomonas maltophilia* R551-3 was used as outgroup to root the tree. Clade credibility values are shown above branches. Species are abbreviated as follows: Xca = *X. campestris*, Xci = *X. citri*, Xci malva. = *X. citri* subsp. *malvacearum*, Xax = *X. axonopodis*, Xfu = *X. fuscans*, Xfu auran. = *X. fuscans* subsp. *aurantifolii*, *X. maculifoliogardeniae* (strain from a Gardenia plant in Virginia, not a validly published species), Xal = *X. alfalfae*, Xor = *X. oryzae*, and Xho = *X. hortorum*. Subspecies and pathovar names are indicated in the figure without the “ssp.” and “pv.” prefix, respectively, to save space. Numbers within parentheses next to strain names indicate the total number of strains with the same sequence in the Plant Associated Microbes Database (PAMDB). Roman numerals indicate groups analyzed in regard to genetic distance in Table 2. Numbers 9.1 through 9.6 are the groups identified by Rademaker and colleagues within *X. axonopodis* (30).

catenated allele sequences for any chosen strain combination (or all strains in the database). Sequences are displayed in a new window in fasta format (with headers displaying strain and locus names). Users have the option to download only nonredundant sequences (unique alleles). In this case, all strains having the same alleles will be listed in the respective sequence header. These data can then be used to build phylogenetic trees and perform molecular evolution and population genetics analysis using third-party software (not included in PAMDB).

Data submission. To submit data to PAMDB, users need to send an e-mail request to the PAMDB administrator to become “submitters”. This allows the administrator to verify the identity of submitters and avoid submission of data sets of low quality by unknown users.

Submitters can add and edit metadata of their strains directly on the website or send tables to the PAMDB administrator for bulk uploading. Because photographic data can be easily saved in PAMDB, we hope that PAMDB will encourage the collection and submission of images of disease symptoms. These data could significantly improve disease diagnostics and submission of images of symptoms to PAMDB will help preserve these data that would otherwise be lost because images are often not published in journals due to printing costs.

To assign allele sequences to strains, submitters compare their sequence data to the alleles in the database using the BLAST program (3) as described above. If submitters find that a strain has a sequence that is 100% identical over the entire length to an allele already in the database, submitters can add that allele number to their strain by themselves or send a table with allele data for all their strains to the PAMDB administrator. The PAMDB administrator may request submitters to provide their raw sequence data to assure the quality of data submitted to PAMDB.

If a strain has an allele sequence not yet in PAMDB, the chromatogram of that allele sequence needs to be sent by e-mail to the PAMDB administrator and, if of good quality, the new allele will be added to PAMDB by the administrator and assigned a new allele number. Submitters can then assign that allele number to their strains by themselves. It is important to note that the PAMDB administrator will not be able to independently verify that the sequence provided came from a particular strain.

Therefore, the PAMDB administrator cannot give any guarantee regarding correctness of allele assignments. Moreover, submitted data may be unpublished data. Users who want to use unpublished data deposited in PAMDB are responsible for contacting submitters to verify quality of these data (submitter e-mails are automatically included for each strain in PAMDB). We also encourage submitters to make strains from which sequences are submitted to PAMDB available to other users on request. For strains with new allele sequences, we encourage strain submission to one or more international culture collections.

A particular feature of PAMDB is that metadata and allele data can be added and appended to records subsequent to record generation, as they become available. For example, data on substrates can be submitted prior to obtaining detailed information on the associated bacteria. This setup of PAMDB provides a flexible infrastructure for storage of information about isolation of bacteria, starting from sample collection all the way to genetic analysis.

A phylogeny of *X. axonopodis* and related strains. Young and co-workers (47) recently published an MLSA study of the genus *Xanthomonas* using four loci previously found to give congruent phylogenies (13), thus providing an excellent example of the power of MLSA in elucidating the phylogeny of an entire plant pathogen group. In addition to adding the MLSA data obtained by Young and colleagues (47) to PAMDB, here we also report a second MLST/MLSA scheme for *Xanthomonas* using six loci. Only one of these loci, *gyrB*, is in common with the MLSA scheme used by Young and colleagues (47). However, even for that locus, a different nonoverlapping gene fragment was used. Loci were chosen among previously developed MLST schemes of *P. syringae* (32) and *Xylella* (33) (Table 1 provides locus details and primers) and without prescreening for congruency to allow for future unbiased determination of recombination between loci. Allele data for 85 strains (e-Xtra Table) were submitted to PAMDB with a focus on *X. axonopodis*, as proposed by Vauterin and colleagues (41). *X. axonopodis* (41) includes pathogen strains from plant species as diverse as citrus, tomato, pepper, cotton, soybean, bean, grape, and alfalfa. Since then, several pathogens attributed to *X. axonopodis* by Vauterin (41) have been described as distinct species based on polyphasic taxonomy, such as the citrus canker pathogens *X. citri* and *X. fuscans* subsp. *aurantifolli*,

TABLE 3. List of tables in the Plant Associated Microbes Database (PAMDB) MySQL database

Table no.	Table name	Table description	Columns ^a
1	Organism	Assigns an identification (ID) and an initial to an organism at the genus level	organism_id (pk, ai), organism_name, initial
2	Species	Assigns an ID to species	species_id (pk, ai), species_name, organism (T 1)
3	Pathovar	Assigns an ID to subspecies, pathovars, or other subspecies category	pathovar_id (pk, ai), pathovar_name, organism (T 1)
4	Strain	Contains metadata of strains	strain_id (pk, ai), genus, species (T 2), pathovar_subspecies (T 3), other_name, plant_scientific, plant_common, cultivar_ecotype, town, region, country, year, latitude, longitude, host_range, strain_comment, type_of_isolation, from_third_party, substrate, plate, colony description, num_of_colonies, unit_type, strain_name, num_ATCC, num_CFBP, num_ICMP, num_NCPPB, num_MAFF, num_LMG, num_ICMP, oxidase, esculin_hydrolysis, arginine_dihydrolase, HR_on_tobacco, INA_temp_10e4, INA_temp_10e6, syringomycin_production, reference, original_collector, image_1, image_2, image_3, organism (T 1)
5	Locus	Assigns locus IDs to loci belonging to MLST schemes for different organisms	locus_id, locus_name, locus_schema, comment, organism (T 1)
6	Strain locus	Assigns allele numbers to loci of strains	strain_id (T 4), locus_id (T 5), allele_number (T 7), username, organism (T 1)
7	Sequence	Assigns DNA sequences to allele numbers	seq_id (pk, ai), seq_num, seq_locus_id (T 5), sequence_nt, organism (T 1)
8	State	Assigns abbreviations to US states	state_id (pk), state, state_abbreviation
9	Country	Assigns codes to country names	country_code (pk), country_name
10	Substrate	Contains metadata of substrates	substrate_id (pk, ai), type, date, country (T 9), state (T 8), city, latitude, longitude, pre_treatment, ecotype_cultivar, plant_scientific, plant_common, notes, username, image_1, image_2, image_3, organism (T 1)
11	User	Contains user data	email (pk), first_name, last_name, city, state (T 8), country (T 9), institution, access_privilege, password

^a Definitions: pk = primary key; ai = autoincrement; T no. = primary key of table number; “other_name” = synonyms and historically published names; and “strain_name” = the words “type strain” or “pathotype strain” follow the strain name when appropriate.

the citrus leaf spot pathogen *X. alfalfae* subsp. *citrumelonis* (34,35), and the tomato and pepper pathogens *X. vesicatoria*, *X. euvesicatoria*, *X. gardneri*, and *X. perforans* (23). We show here how the MLSA data in PAMDB can be used to assess the relationships among these strains and more distantly related xanthomonads. To do this, the concatenated data set of the six loci was downloaded for all strains from PAMDB. A phylogenetic tree was constructed from these data with the program MrBayes (19,31) using *S. maltophilia* as outgroup (Fig. 1). A neighbor-joining tree constructed with the same data and that has the same overall topology confirming the relationships between strains inferred with the Bayesian approach is shown in the online e-Xtra. Most branches have posterior inclusion probabilities between 0.9 and 1. Although all gene fragments were different from those used in the study by Young and co-workers (47), the pathogens that were used in both studies group with high statistical support into very similar clades, confirming the results obtained by Young and co-workers (47). The identified clades also largely correspond to the groups previously identified by Rademaker and colleagues using BOX, ERIC, and REP PCR (30).

Comparison of diversity existing within *P. syringae*, *Xanthomonas* spp., and *R. solanacearum*. MLSA data for *Xanthomonas* spp., *P. syringae*, and *R. solanacearum* deposited in PAMDB make it easy to compare population genetic parameters among pathogens. To demonstrate how PAMDB can be used in this regard, we decided to determine the diversity that exists within each of these three important plant pathogen groups. Mean genetic distances were calculated using two different methods using concatenated loci for each group. For *Xanthomonas* spp., we also considered the following phylogenetic subgroups separately (Fig. 1): phylogroup I, corresponding to *X. campestris* (41); phylogroup II, corresponding to *X. axonopodis* (41); phylogroup IIa, corresponding approximately to *X. citri* (34,35); phylogroup IIb, corresponding approximately to *X. fuscans* (34,35); and phylogroup IIc, corresponding approximately to *X. alfalfae* (34,35). For *P. syringae*, we considered phylogroups I, II, and III (20) (as labeled in the online e-Xtra) corresponding approximately to genomospecies 3, 1, and 2 (15), respectively. For *R. solanacearum*, we considered phylogroups I, II, and IV (8) (as labeled in the online e-Xtra). We also identified the pair of strains in each group with maximum genetic distance by simply looking for the pair of strains that were most different from each other based on percentage of nucleotide differences. Mean genetic distance and maximum genetic distance correlated well in all groups (Table 2). *P. syringae* appears to be the most diverse group, with mean and maximum genetic distance values greater than even the respective values for the entire *Xanthomonas* genus. In contrast, *R. solanacearum* had the lowest diversity of the three groups. Within *P. syringae*, the three phylogroups each had diversity values similar to *X. axonopodis* (41), ranging between those of *R. solanacearum* overall and *R. solanacearum* phylogroup II. Among *Xanthomonas* spp., *X. axonopodis* (phylogroup II) has almost the same diversity as the entire *R. solanacearum* complex. The lowest diversity of all groups is displayed by *R. solanacearum* phylogroups I and IV and *Xanthomonas* phylogroups I, IIa, IIb, and IIc.

Examples how MLSA can distinguish between pathogens causing diseases with almost identical symptoms. Citrus canker, a devastating disease of citrus trees, can be caused by two different organisms, *X. citri* subsp. *citri* and *X. fuscans* subsp. *aurantifolii* (35). Within the latter subspecies, two groups named B and C can be further distinguished. Although these organisms cause identical disease symptoms, they have different host ranges (38). Their geographic distribution is also different. Both *X. fuscans* subsp. *aurantifolii* and *X. citri* subsp. *citri* are present in South America whereas *X. citri* subsp. *citri* is the only citrus canker pathogen present in Florida (6). It is important to dis-

tinguish between these organisms to be able to monitor their dissemination with the goal of preventing their spread to new geographic regions. Based on the data accessible in PAMDB, MLST can easily distinguish between *X. citri* subsp. *citri* and *X. fuscans* subsp. *aurantifolii*: five of six MLST loci distinguish between the typical *X. citri* subsp. *citri* strains from Florida and the *X. fuscans* subsp. *aurantifolii* group B strain 1622 and all six loci distinguish between *X. citri* subsp. *citri* strains from Florida and the *X. fuscans* subsp. *aurantifolii* group C strain ATCC51302. Two loci distinguish between *X. fuscans* subsp. *aurantifolii* group B and group C strains.

In recent years, *P. syringae* pv. *alisalensis* caused severe economic losses on cruciferous plants from the Brassicaceae family across the United States (C. T. Bull, *personal communication*). Although this organism has a distinct host range from strains classified as *P. syringae* pv. *maculicola* in regard to pathogenicity to the cruciferous plant broccoli raab and oat (9), symptoms caused by these two organisms are almost indistinguishable from each other on other plants of the Brassicaceae family. To determine the epidemiology of *P. syringae* pv. *alisalensis* and to limit its spread, it is important to distinguish between *P. syringae* pv. *alisalensis* and *P. syringae* pv. *maculicola*. Using data deposited in PAMDB, the pathotype strain *P. syringae* pv. *alisalensis* BS91 is different in all five sequenced MLST loci from the pathotype strain of *P. syringae* pv. *maculicola* LMG 5071, making it rapid and easy to distinguish between the two by sequencing any one of these loci.

DISCUSSION

PAMDB is the first MLST/MLSA database for plant-pathogenic and plant-associated bacteria. It was necessary to design a database specifically for plant-associated bacteria because the existing MLST database software (22) was designed for human pathogens and does not provide the possibility to add data that are crucial for plant pathogens; in particular, and most importantly, images of disease symptoms. Moreover, MLST databases for human pathogens are curated: data are checked for quality and added to the databases by specialized personnel. This is clearly an advantage in regard to reliability of data. However, it would be cost prohibitive to maintain curated databases for many plant-pathogenic bacteria. Therefore, PAMDB was designed to make it easy for users to add data to the database themselves. However, submitter status is moderated by the site administrators to assure that only qualified users will submit data to PAMDB.

Data have already been entered in PAMDB for four pathogen groups, greatly facilitating the accessibility and use of previously published data for *A. citrulli* and *A. avenae* (14), *P. syringae* (28,32,45), *R. solanacearum* (8), and *Xanthomonas* spp. (47). A new, previously unpublished, MLST/MLSA scheme for *Xanthomonas* spp. is also provided. Importantly, the scalable design of the PAMDB database and website provides the flexibility to easily add additional species and typing schemes as they become available. For example, based on whole-genome comparisons, we have identified loci that can distinguish between *P. syringae* strains isolated from tomato that are identical in the loci currently in PAMDB (B. A. Vinatzer, *unpublished data*) and we will add these additional loci in the future, which will further increase the resolution of subtyping *P. syringae* pv. *tomato*.

An example of the power of PAMDB is the phylogenetic tree shown in Figure 1, built with the concatenated MLSA data for the genus *Xanthomonas*. Most branches of the tree have posterior inclusion probabilities between 0.9 and 1. Therefore, new strains belonging to any of the phylogenetic groups present in the tree can be assigned to these groups using MLSA with confidence. However, MLSA usually has only limited power at distinguishing between strains belonging to the same pathovar or subspecies. Although we showed that the two citrus canker pathogens *X.*

fuscans subsp. *aurantifolii* and *X. citri* subsp. *citri* are easily distinguishable, and *P. syringae* pv. *maculicola* can be easily distinguished from *P. cannabina* pv. *alisalensis*, most *X. citri* subsp. *citri* strains from Florida were identical in all six loci to the sequenced *X. citri* subsp. *citri* strain 306 from Brazil (10). However, strains of *X. citri* subsp. *citri* Wellington (A^W) from Florida and strains of *X. citri* subsp. *citri* A* from southwest Asia, previously found to have different host ranges compared with the most common *X. citri* subsp. *citri* strains in Florida (2), were confirmed to be different from all other *X. citri* subsp. *citri* strains in one MLSA locus. Therefore, MLSA has some power even to distinguish between strains within *X. citri* subsp. *citri*, allowing an initial subtyping of this organism. Whole-genome sequencing of representative *X. citri* subsp. *citri* strains should allow for the selection of additional informative MLST loci to subtype *X. citri* subsp. *citri* in order to investigate the epidemiology of citrus canker and the microevolution of this devastating pathogen.

The comparison of the diversity existing within *P. syringae*, *Xanthomonas* spp., and *R. solanacearum* confirmed that strains from *P. syringae* sensu lato comprise a very diverse group of pathogens; even more diverse than the entire genus of *Xanthomonas*. This underlines the importance of defining taxonomic groups within *P. syringae*, especially because strains with the same pathovar designation are found in two to three different phylogroups, indicating that they may be misidentified. The data presented here indicate that MLST could be used to develop proposals that should allow for the genomospecies identified by Gardan and colleagues (15) to be readily distinguished from one another and be elevated to species. In the meantime, PAMDB makes it possible to precisely attribute unknown *P. syringae* strains to a defined phylogenetic group even in the absence of adequate taxonomy. To further facilitate identification of *P. syringae* strains using MLST, we are currently analyzing most described pathotype strains of *P. syringae* to include them in PAMDB.

In the genus *Xanthomonas*, our analysis confirmed the high diversity existing in *X. axonopodis* (41) compared with *X. campestris* and showed that the recently described species *X. citri*, *X. fuscans*, and *X. alfalfae* (35) each have diversity values that are similar to those of *X. campestris* (41). However, because these species designations have not yet been widely used by the *Xanthomonas* research community and other classifications are also in use, MLSA has an important role in attributing *Xanthomonas* strains to phylogenetic groups, regardless of the names used for these groups.

The *R. solanacearum* species complex is significantly less diverse than the *P. syringae* species complex and the *Xanthomonas* genus. Although a description of species within *R. solanacearum* may be useful, the current practice of attributing strains to phylogroups (4) fulfills the needs of the *R. solanacearum* community. Moreover, as with *P. syringae* and *Xanthomonas* spp., MLSA also offers the possibility of precisely attributing strains to defined phylogenetic subgroups within phylogroups independently of the current regulated nomenclature.

With the continuing reduction in costs for whole-genome sequencing, complete genome sequences of many bacteria deposited in PAMDB will become available. Instead of sequencing five or six loci of a strain, in a few years it may become economically feasible to sequence the entire genome of an unknown strain to sufficient coverage to compare it with reference genomes. Although PAMDB is not currently able to handle whole-genome data, we envision future development of PAMDB to allow deposition and comparison of whole-genome data. Although new tables and new websites will be required, it will be possible to preserve all the strain information currently in PAMDB.

In conclusion, our vision for PAMDB is for it to grow into an ever-expanding resource for plant disease diagnostics and the study of epidemiology, population genetics, and molecular evolution of plant-associated bacteria with a steadily increasing number

of strains from a growing number of organisms, and with sequence data that will include entire genomes.

ACKNOWLEDGMENTS

Research in the Vinatzer lab was funded by the NSF (Award IOS 0746501), Virginia Tech start-up funds, and a Virginia Tech ASPIRES grant. N. F. Almeida was partly supported by the CNPq grant 200447/2007-6. We thank the Virginia Tech undergraduate students T. Wiseman, N. Long, and L. Williams for their help with DNA extractions, PCR, and DNA sequence analysis.

LITERATURE CITED

- Ah-You, N., Gagnevin, L., Grimont, P. A., Brisse, S., Nesme, X., Chireleu, F., Bui Thi Ngoc, L., Jouen, E., Lefeuvre, P., Verniere, C., and Pruvost, O. 2009. Polyphasic characterization of xanthomonads pathogenic to members of the *Anacardiaceae* and their relatedness to species of *Xanthomonas*. *Int. J. Syst. Evol. Microbiol.* 59:306-318.
- Al-Saadi, A., Reddy, J. D., Duan, Y. P., Brunings, A. M., Yuan, Q., and Gabriel, D. W. 2007. All five host-range variants of *Xanthomonas citri* carry one *pthA* homolog with 17.5 repeats that determines pathogenicity on citrus, but none determine host-range variation. *Mol. Plant-Microbe Interact.* 20:934-943.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
- Alvarez, A. M. 2005. Diversity and diagnosis of *Ralstonia solanacearum*. Pages 437-447 in: *Bacterial Wilt, the Disease and the Ralstonia solanacearum Species Complex*. C. Allen, P. Prior, and A. C. Hayward, eds. American Phytopathological Society, St. Paul, MN.
- Brehony, C., Jolley, K. A., and Maiden, M. C. 2007. Multilocus sequence typing for global surveillance of meningococcal disease. *FEMS Microbiol. Rev.* 31:15-26.
- Brunings, A. M., and Gabriel, D. W. 2003. *Xanthomonas citri*: breaking the surface. *Mol. Plant Pathol.* 4:141-157.
- Bull, C. T., De Boer, S. H., Denny, T. P., Firrao, G., Fischer-Le Saux, M., Sandler, G. S., Scortichini, M., Stead, D. E., and Takikawa, Y. 2008. Demystifying the nomenclature of bacterial plant pathogens. *J. Plant Pathol.* 90:403-417.
- Castillo, J. A., and Greenberg, J. T. 2006. Evolutionary dynamics of *Ralstonia solanacearum*. *Appl. Environ. Microbiol.* 73:1225-1238.
- Cintas, N. A., Koike, S. T., and Bull, C. T. 2002. A new pathovar, *Pseudomonas syringae* pv. *alisalensis* pv. nov., proposed for the causal agent of bacterial blight of broccoli and broccoli raab. *Plant Dis.* 86:992-998.
- da Silva, A. C., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., Quaggio, R. B., Monteiro-Vitorello, C. B., Van Sluys, M. A., Almeida, N. F., Alves, L. M., do Amaral, A. M., Bertolini, M. C., Camargo, L. E., Camarotte, G., Cannavan, F., Cardozo, J., Chamberg, F., Ciapina, L. P., Cicarelli, R. M., Coutinho, L. L., Cursino-Santos, J. R., El-Dorry, H., Faria, J. B., Ferreira, A. J., Ferreira, R. C., Ferro, M. I., Formighieri, E. F., Franco, M. C., Greggio, C. C., Gruber, A., Katsuyama, A. M., Kishi, L. T., Leite, R. P., Lemos, E. G., Lemos, M. V., Locali, E. C., Machado, M. A., Madeira, A. M., Martinez-Rossi, N. M., Martins, E. C., Meidanis, J., Menck, C. F., Miyaki, C. Y., Moon, D. H., Moreira, L. M., Novo, M. T., Okura, V. K., Oliveira, M. C., Oliveira, V. R., Pereira, H. A., Rossi, A., Sena, J. A., Silva, C., de Souza, R. F., Spinola, L. A., Takita, M. A., Tamura, R. E., Teixeira, E. C., Tezza, R. I., Trindade dos Santos, M., Truffi, D., Tsai, S. M., White, F. F., Setubal, J. C., and Kitajima, J. P. 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417:459-463.
- DeLey, J. 1978. Modern molecular methods in bacterial taxonomy: evaluation, application, prospects. Pages 347-357 in: 4th Int. Conf. Plant Pathogenic Bacteria. Angers, France.
- Dye, D. W., Bradbury, J. F., Goto, M., Hayward, A. C., Lelliott, R. A., and Schroth, M. N. 1980. International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. *Rev. Plant Pathol.* 142:153-158.
- Fargier, E., and Manceau, C. 2006. Pathogenic and genetic characterization of *Xanthomonas campestris* by plant inoculations and multilocus site analysis. Page 35 in: 11th Int. Conf. Plant Pathogenic Bacteria. J. Elphinstone, R. Weller, R. Thwaites, N. Parkinson, D. Stead, and G. Sandler, eds. Edinburgh.
- Feng, J., Schuenzel, E. L., Li, J., and Schaad, N. W. 2009. Multilocus sequence typing reveals two evolutionary lineages of *Acidovorax avenae* subsp. *citrulli*. *Phytopathology* 99:913-920.

15. Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., and Grimont, P. A. 1999. DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (ex Sutic and Dowson 1959). *Int. J. Syst. Bacteriol.* 49:469-478.
16. Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F. L., and Swings, J. 2005. Opinion: Re-evaluating prokaryotic species. *Nat. Rev. Microbiol.* 3:733-739.
17. Goss, E. M., Kreitman, M., and Bergelson, J. 2005. Genetic diversity, recombination and cryptic clades in *Pseudomonas viridiflava* infecting natural populations of *Arabidopsis thaliana*. *Genetics* 169:21-35.
18. Hauben, L., Vauterin, L., Swings, J., and Moore, E. R. 1997. Comparison of 16S ribosomal DNA sequences of all *Xanthomonas* species. *Int. J. Syst. Bacteriol.* 47:328-335.
19. Huelsenbeck, J. P., and Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.
20. Hwang, M. S., Morgan, R. L., Sarkar, S. F., Wang, P. W., and Guttman, D. S. 2005. Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. *Appl. Environ. Microbiol.* 71:5182-5191.
21. Johnson, J. L. 1984. Nucleic acids in bacterial classification. Pages 8-11 in: *Bergey's Manual of Systematic Bacteriology*. N. R. Krieg and J. G. Holt, eds. Williams and Wilkins, Baltimore, MD.
22. Jolley, K. A., Chan, M. S., and Maiden, M. C. 2004. mlstdbNet—distributed multi-locus sequence typing (MLST) databases. *BMC Bioinf.* 5:86.
23. Jones, J. B., Lacy, G. H., Bouzar, H., Stall, R. E., and Schaad, N. W. 2004. Reclassification of the xanthomonads associated with bacterial spot disease of tomato and pepper. *Syst. Appl. Microbiol.* 27:755-762.
24. Lapage, S. P., Sneath, P. H. A., F. L. E., Skerman, V. B. D., Seeliger, H. P. R., and Clark, W. A. 1975. *International Code of Nomenclature of Bacteria*. American Society for Microbiology, Washington, DC.
25. Lee, B. M., Park, Y. J., Park, D. S., Kang, H. W., Kim, J. G., Song, E. S., Park, I. C., Yoon, U. H., Hahn, J. H., Koo, B. S., Lee, G. B., Kim, H., Park, H. S., Yoon, K. O., Kim, J. H., Jung, C. H., Koh, N. H., Seo, J. S., and Go, S. J. 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Res.* 33:577-586.
26. Maiden, M. C. 2006. Multilocus sequence typing of bacteria. *Annu. Rev. Microbiol.* 26:561-588.
27. Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA* 95:3140-3145.
28. Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffiere, A., Yan, S., Dominguez, H., and Thompson, B. M. 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *ISME J.* 2:321-334.
29. Perez-Losada, M., Browne, E. B., Madsen, A., Wirth, T., Viscidi, R. P., and Crandall, K. A. 2006. Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. *Infect. Genet. Evol.* 6:97-112.
30. Rademaker, J. L., Louws, F. J., Schultz, M. H., Rossbach, U., Vauterin, L., Swings, J., and de Bruijn, F. J. 2005. A comprehensive species to strain taxonomic framework for *Xanthomonas*. *Phytopathology* 95:1098-1111.
31. Ronquist, F., and Huelsenbeck, J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
32. Sarkar, S. F., and Guttman, D. S. 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl. Environ. Microbiol.* 70:1999-2012.
33. Scally, M., Schuenzel, E. L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Appl. Environ. Microbiol.* 71:8491-8499.
34. Schaad, N. W., Postnikova, E., Lacy, G., Sechler, A., Agarkova, I., Stromberg, P. E., Stromberg, V. K., and Vidaver, A. K. 2006. Emended classification of xanthomonad pathogens on citrus. *Syst. Appl. Microbiol.* 29:690-695.
35. Schaad, N. W., Postnikova, E., Lacy, G. H., Sechler, A., Agarkova, I., Stromberg, P. E., Stromberg, V. K., and Vidaver, A. K. 2005. Reclassification of *Xanthomonas campestris* pv. *citri* (ex Hasse 1915) Dye 1978 forms A, B/C/D, and E as *X. smithii* subsp. *citri* (ex Hasse 1915) sp. nov. nom. rev. comb. nov., *X. fuscans* subsp. *aurantifolii* (ex Gabriel 1989) sp. nov. nom. rev. comb. nov., and *X. alfalfae* subsp. *citrumelo* (ex Riker and Jones) Gabriel et al., 1989 sp. nov. nom. rev. comb. nov.; *X. campestris* pv. *malvacearum* (ex smith 1901) Dye 1978 as *X. smithii* subsp. *smithii* nov. comb. nov. nom. nov.; *X. campestris* pv. *alfalfae* (ex Riker and Jones, 1935) dye 1978 as *X. alfalfae* subsp. *alfalfae* (ex Riker et al., 1935) sp. nov. nom. rev.; and "var. *fuscans*" of *X. campestris* pv. *phaseoli* (ex Smith, 1987) Dye 1978 as *X. fuscans* subsp. *fuscans* sp. nov. *Syst. Appl. Microbiol.* 28:494-518.
36. Skerman, V. B. D., McGowan, V., and Sneath, P. H. A. 1980. Approved lists of bacterial names. *Int. J. Syst. Bacteriol.* 30:225-420.
37. Stackebrandt, E., and Goebel, B. M. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* 44:846-849.
38. Stall, R. E., and Seymour, C. P. 1983. Canker, a threat to citrus in the Gulf coast states. *Plant Dis.* 67:581-585.
39. Swofford, D. L. 2003. PAUP* Phylogenetic Analysis Using Parsimony. Sinauer Associates, Sunderland, MA.
40. Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596-1599.
41. Vauterin, L., Hoste, B., Kersters, K., and Swings, J. 1995. Reclassification of *Xanthomonas*. *Int. J. Syst. Bacteriol.* 45:472-489.
42. Vinatzer, B. A., and Bull, C. T. 2009. The impact of genomics approaches on our understanding of diversity and taxonomy of plant pathogenic bacteria. Pages 37-61 in: *Plant Pathogenic Bacteria: Genomics and Molecular Biology*. R. W. Jackson, ed. Horizon Scientific Press, Norfolk, UK.
43. Wang, P. W., Morgan, R. L., Scortichini, M., and Guttman, D. S. 2007. Convergent evolution of phytopathogenic pseudomonads onto hazelnut. *Microbiology* 153:2067-2073.
44. Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, L., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P., and Truper, H. G. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* 37:463-464.
45. Yan, S., Liu, H., Mohr, T. J., Jenrette, J., Chiodini, R., Zaccardelli, M., Setubal, J. C., and Vinatzer, B. A. 2008. Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000, a very atypical tomato strain. *Appl. Environ. Microbiol.* 74:3171-3181.
46. Young, J. M., Dye, D. W., and Wilkie, J. P. 1978. A proposed Nomenclature and classification for plant pathogenic bacteria. *N. Z. J. Agric. Res.* 21:158-161.
47. Young, J. M., Park, D. C., Shearman, H. M., and Fargier, E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. *Syst. Appl. Microbiol.* 31:366-377.
48. Young, J. M., Saddler, G., Takikawa, Y., De Boer, S. H., Vauterin, L., Gardan, L., Gvozdyak, R. I., and Stead, D. E. 1996. Names of plant pathogenic bacteria 1864-1995. *Rev. Plant Pathol.* 75:721-763.