Amplified fragment length polymorphism and multilocus sequence analysis-based genotypic relatedness among pathogenic variants of *Xanthomonas citri* pv. *citri* and *Xanthomonas campestris* pv. *bilvae*

Lan Bui Thi Ngoc, Christian Vernière, Emmanuel Jouen, Nathalie Ah-You, Pierre Lefeuvre, Frédéric Chiroleu, Lionel Gagnevin and Olivier Pruvost

INTRODUCTION

The classification of some phytopathogenic bacterial species was initially based on host specialization, although this was evidently not sound taxonomically (Dowson, 1939; Stolp et al., 1965). Nevertheless, plant pathologists and phytosanitary policy regulators required names for plant-pathogenic bacteria that clearly reflected their pathogenic ability. This led to the development of the pathovar classification at an infrasubspecific level (Dye et al., 1980). Furthermore, strains with differential host ranges within a pathovar were sometimes described (e.g. xanthomonads pathogenic to citrus or cassava) and often referred to as pathotypes (Civerolo & Fan, 1982; Verdier et al., 1998).

A total of four xanthomonads are pathogenic to plant species within the Rutaceae. Two pathogens that cause spot diseases of citrus (i.e. water-soaked spots turning into flat, necrotic lesions) have been described as two distinct *Xanthomonas* pathovars, namely *Xanthomonas axonopodis* pv. *citrumelo* (Vauterin et al., 1995) (synonyms *Xanthomonas alfalfae* pv. *citrumelo* and *Xanthomonas campestris* pv. *citri* pathotype E), the causal agent of citrus bacterial spot, and *X. campestris* pv. *bilvae*, a pathogen that causes spots on rutaceous plants in India (Patel et al., 1953), which was not
reclassified by Vauterin et al. (1995) or in more recent studies. In addition, two pathogens that cause canker-like diseases of citrus have also been described as distinct pathovars. All bacterial pathogens that cause canker-like diseases of citrus are regarded as quarantine organisms in many countries, and high-cost, drastic control measures are often implemented. Xanthomonas citri pv. citri (syonyms X. axonopodis pv. citri and X. campestris pv. citri pathotype A) is the causal agent of Asiatic canker (Ah-You et al., 1998; Schaad et al., 2006). This pathogen induces erumpent, callus-like lesions with a water-soaked margin on leaf, fruit and stem tissue. Severe attacks cause extensive defoliation, premature fruit drop and twig dieback. The hypothesis has been proposed that citrus canker originated in India (in the 1830s) or Java (in the 1840s), rather than other regions, because the oldest herbarium specimens with canker lesions originated from these areas (Fawcett & Jenkins, 1933). Citrus canker now occurs in more than 30 countries throughout the world, including in Africa, where it is currently re-emerging (Balestra et al., 2008; Traoré et al., 2008). Its presence in many areas is a continuous threat to citriculture (Civerolo, 1984). This pathogen is of major concern because of its wide host range among citrus and related species, extreme aggressiveness and widespread distribution worldwide (Gottwald et al., 2002). Strains of this pathogen will be referred to as pathotype A strains hereafter. X. citri pv. aurantifolii (syonyms Xanthomonas fuscans pv. aurantifolii, X. axonopodis pv. aurantifolii and X. campestris pv. citri pathotypes B and C) is responsible for the milder, geographically restricted South American canker, pathogenic to a narrower host range, including lemons and limes, which are the most susceptible species (Ah-You et al., 2009; Schaad et al., 2006). In comparison with X. citri pv. citri, X. citri pv. aurantifolii has a lower economic incidence on citrus production. These two pathogens are phenotypically and genetically distinct (Hartung & Civerolo, 1989; Pruvost et al., 1992; Rademaker et al., 2005; Vernière et al., 1993). More recently, xanthomonads with a narrow host range but shown to be genetically related to X. citri pv. citri using RFLP analysis were reported from different areas of south-west and central Asia, including Saudi Arabia, Oman, Iran, India, Thailand and Cambodia (Bui Thi Ngoc et al., 2007, 2008; Vernière et al., 1998). These strains, designated pathotype A*, all originated from Mexican lime (Citrus aurantifolia) and produced canker-like lesions on this host species after inoculation, but not on grapefruit (Citrus paradisi), a species highly susceptible to X. citri pv. citri (Vernière et al., 1998). Finally, strains naturally pathogenic to Mexican lime and alemow (Citrus macrophylla) were detected in Florida and designated pathotype A# (Sun et al., 2004). These strains did not produce an ampiclon when assayed using specific primers designed for pathotype A strains (Hartung et al., 1993; Sun et al., 2004). Recent DNA–DNA hybridization data suggested that pathotype A* and A# strains should be classified within the X. citri genospecies (Sun et al., 2004). However, no large collection of strains has been studied using high-throughput techniques, as recommended by Stackebrandt et al. (2002).

Amplified fragment length polymorphism (AFLP) and multilocus sequence analysis (MLSA) are used here to gain insight into phylogenetic relatedness among X. citri pv. citri strains and relatives. AFLP (Vos et al., 1995) is based on arbitrary genome sampling and generates large numbers of markers over the whole genome. AFLP has proven useful for resolving the species status of several xanthomonads (Ah-You et al., 2009; Janssen et al., 1996; Rademaker et al., 2000; Roumagnac et al., 2004), especially after calculating evolutionary genome divergence (EGD) between individuals. EGD estimates the number of nucleotide substitutions that have occurred since the strains diverged from a common ancestor, and is well correlated with other evolutionary indicators (Mougel et al., 2002). Recently, MLSA, based on sequence analysis of several housekeeping genes, has been developed for species delineation (Chelo et al., 2007; Gevers et al., 2005; Hanage et al., 2005a, b; Martens et al., 2007; Richter et al., 2006), including members of Xanthomonas. The advantage of MLSA is that phylogenetic relationships of large sets of strains can be analysed with a better portability than for genotyping techniques such as AFLP or rep-PCR.

We have established previously (Ah-You et al., 2009) that AFLP and MLSA based on atpD (encoding the F1-F0 ATPase β subunit), dnaK (70 kDa heat-shock protein, Hsp70) and gyrB (DNA gyrase subunit β) sequences have the potential to discriminate among Xanthomonas species, and the latter study yielded a redefinition of Xanthomonas citri. The present study is a follow-up that aimed to complement our MLSA scheme with an additional partial gene sequence, efp (encoding elongation factor P), and to evaluate the pathological and phylogenetic relatedness to different Xanthomonas species and pathovars of a large set of strains pathogenic to rutaceous host species (including the unreclassified citrus pathogen X. campestris pv. bilvae).

METHODS

Bacterial strains and DNA extraction. Bacterial strains used in this study are listed in Supplementary Table S1, available in IJSEM Online. Cultures were stored as lyophiles and/or at −80 °C in microbank tubes (Fisher Scientific). Cultures were obtained on YPGA (1 l–1: 7 g yeast extract, 7 g peptone, 7 g glucose, 18 g agar, 20 mg propiconazole, pH 7.2), incubated at 28 °C for 3–8 days. Single colonies were subcultured on YPGA for 24 h at 28 °C. These subcultures were used to inoculate 4 ml YP broth tubes, which were incubated at 28 °C under agitation for 16–18 h. These suspensions were used for DNA extraction using the DNeasy tissue kit (Qiagen) following the manufacturer’s instructions and DNA concentrations were estimated by fluorimetry (TKO 100 fluorimeter; Hoefer).

Pathogenicity tests. Strains submitted to pathogenicity tests are shown in Supplementary Table S1. They included strains of pathovars aurantifolii and bilvae and pathogenicity variants of pathovar citri (i.e. pathotypes A*, A# and A*). Inoculations, as described previously (Vernière et al., 1998), were performed on five Citrus species: C. sinensis (pineapple sweet orange), C. aurantifolia (Mexican lime), C. macrophylla (alemow), C. latifolia (Tahiti lime) and C. paradisi (Marsh grapefruit).Immature fully expanded citrus leaves were sterilized by soaking for 2 min in 1% sodium hypochlorite and rinsed three times in sterile distilled water.
Leaves were placed on the surface of 1% water agar plates with adaxial surfaces facing upwards. Ten wounds per leaf were made with a sterile needle and droplets (5 μl) of bacterial suspensions containing approximately 1 x 10⁶ cfu ml⁻¹ were placed on each wound. Bacterial suspensions used for inoculation were prepared in 0.01 M sterile Sigma 7-9 Tris buffer (pH 7.2) (Sigma-Aldrich) and obtained by tenfold dilutions of spectrophotometrically adjusted suspensions (an OD₆₀₀ of 0.05 is approximately equivalent to 10⁶ cfu ml⁻¹) prepared from 18 h old cultures on YPGA. Sterile Tris buffer was used as inoculum for negative controls. Leaves were incubated in a growth chamber at 28 °C with a photoperiod of 12 h and observed for the development of tissue hyperplasia after 7–21 days. For pv. bilvae strains, attached leaves of C. sinensis, C. aurantifolia, C. macrophylla and C. paradisi were inoculated by infiltration of bacterial suspensions containing approximately 1 x 10⁶ cfu ml⁻¹ as described previously (Ah-You et al., 2007). Leaves were checked daily for lesion development over 3 weeks and population sizes were determined 1 month after inoculation on KC semi-selective medium as described previously (Ah-You et al., 2007).

In planta population sizes were also determined on 1-year-old potted Marsh grapefruit and alemow plants that were pruned to produce uniformly aged shoots. Nine strains were used for inoculation (Supplementary Table S1). The strain collection consisted of four A⁺ strains (originating from Iran, Pakistan, Saudi Arabia and Thailand), one A⁻ strain (from Florida) and four A strains (from Bangladesh, New Zealand, Pakistan and the Philippines). Bacterial suspensions used for inoculation contained approximately 1 x 10⁶ cfu ml⁻¹ and were prepared as explained above. Suspensions were infiltrated (0.4–0.5 cm²) into the mesophyll of mature leaves from the youngest vegetative flush using a sterile syringe. For each strain, 10 inoculation points were performed on each of 12 leaves on different plants. Each experiment was replicated once. Plants were placed in growth chambers at 30 ± 1 °C day and 26 ± 1 °C night and 95 ± 5 % relative humidity for 35 days with a photoperiod of 12 h. Inoculated leaf fragments were sampled 1, 2, 5, 8, 12, 20, 28 and 35 days post-inoculation (p.i.). At each sampling time, ten leaf discs (about 1 cm²) per strain and per host species were homogenized in 5 ml 0.01 M sterile Sigma 7-9 Tris buffer (pH 7.2) (Sigma-Aldrich) using an Ultraturax T25 homogenizer (Janke & Kunkel). Aliquots of the homogenized suspensions and of the tenfold dilutions obtained from the suspensions were plated on KC semi-selective medium (Pruvost et al., 2005) using the Spiral System device (Interscience). Bacterial population sizes were based on enumeration of Xanthomonas-like colonies (starch hydrolysis was checked on doubtful colonies). Data were analysed by ANOVA using R software (version 2.6.1; R Development Core team). Whenever F values were significant (P<0.05), Tukey’s contrasts were used to differentiate amongst the means of the population sizes at a level of 0.05.

PCR assay. A specific nested-PCR assay for X. citri pv. citri was performed using previously developed primers (Hartung et al., 1993, 1996). The assayed strains included all pathotype A⁺ and A⁻ strains, in addition to LMG 9322 and CFBP 2525 reference A strains. Genomic DNA and X. citri pv. citri suspensions in sterile HPLC-grade water that were boiled for 1 min and then cooled on ice were used as templates for PCR. All tests were performed with a GeneAmp PCR system 9700 thermocycler (Applied Biosystems). The first round of PCR was performed with primers 4 (5'-TGTCGTCGCTTGTATGGCAAG-3') and 7 (5'-AGGGTGCGACCCTGCAGGA-3'), which produced a 468 nt fragment for pathotype A strains. The amplification program consisted of 35 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s and extension at 72 °C for 30 s. The primers used in the second PCR round were 94-3 (5'-CGATCCGAGTGTCGATCCTCC-3') and 94-4 (5'-CAGATCGATCGATGAGCGAGGAG-3'), which yielded a 315 nt fragment for pathotype A strains. For this second round of PCR, 1 μl of the first reaction was used as a template. The amplification program consisted of 20 cycles of denaturation at 95 °C for 70 s, annealing at 58 °C for 1 min and extension at 72 °C for 1 min. PCRs were performed in 25 μl reaction mixtures containing 3 mM MgCl₂, 125 μM of each dNTP, 0.5 μM of each primer, 1 μl template DNA and 1.25 U Goldstar Red Taq polymerase (Eurogentec) in 75 mM Tris/HCl, 20 mM (NH₄)₂SO₄, 0.01 % Tween 20 buffer (pH 8.8). PCR products were separated by electrophoresis in 1.5% Seakem LE agarose (FMC Bioproducts), stained with ethidium bromide and visualized with UV light. The amplicons produced from the second round of PCR from two strains (NCPPB 3608 and X2002-1035G) for which no signal was observed after the first PCR round were sequenced (Macrogen).

AFLP analysis. The experiments were performed in 96-well plates in a GeneAmp PCR system 9700 thermocycler (Applied Biosystems), as described previously (Ah-You et al., 2007). Digestions were carried out in a 25 μl volume for 1 h at 37 °C and contained 100 ng bacterial genomic DNA, 10 U Sacl, 2 UMspI (New England Biolabs) and 1 x BSA in 1 x reaction buffer NEB 1. Next, an aliquot of 2.5 μl of the digested products was added to 22.5 μl ligation mixture containing 2 μM Mspl adaptor (Supplementary Table S2), 0.2 μM Sacl adaptor (Supplementary Table S2) (Applied Biosystems), and 2 U T4 DNA ligase (New England Biolabs) in 1 x T4 DNA ligation buffer. Ligation reactions were performed for 3 h at 37 °C before enzyme inactivation at 65 °C for 10 min. Ligation products were diluted tenfold with HPLC-grade water before preselective PCR. The reactions were carried out in 15 μl and contained 5 μl diluted ligation product, 2.5 mM MgCl₂, 0.23 mM each of Mspl and Sacl primers (Supplementary Table S2), 0.45 mM of each dNTP (New England Biolabs) and 0.5 U Goldstar Red Taq DNA polymerase (Eurogentec) in 1 x Goldstar buffer. The following PCR conditions were used: initial extension to ligate the second strand of the adaptors at 72 °C for 2 min, a denaturation step at 94 °C for 2 min, 25 cycles of 94 °C for 30 s, 56 °C for 30 s and 72 °C for 2 min, followed by a final extension step at 72 °C for 10 min. PCR products were diluted tenfold with HPLC-grade water before selective amplification. Selective amplifications using the unlabelled Mspl + A, C, T or G primer and the labelled Sacl + C primer (four different fluorochromes) (Supplementary Table S2) were performed under the same conditions as the preselective PCR, except that the Sacl + C primer concentration was 0.12 μM. The following PCR conditions were used: initial denaturation at 94 °C for 2 min followed by 37 cycles of 94 °C for 30 s, annealing for 30 s at 65 °C for the first cycle, decreased by 0.7 °C per cycle for the next 12 cycles and then 56 °C for the last 24 cycles followed by extension at 72 °C for 2 min with a final extension step at 72 °C for 10 min. Samples were then prepared for capillary electrophoresis by adding 1 μl of the final PCR product to 18.7 μl formamide and 0.3 μl LIZ500 DNA ladder (Applied Biosystems) as an internal standard. The mixture was then denatured for 5 min at 95 °C and placed on ice for at least 5 min. Electrophoresis was performed in an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) using performance-optimized polymer POP-4 at 15 000 V for about 20 min at 60 °C, with an initial injection of 66 s. AFLP fingerprints were analysed visually using GENESCAN software 3.7 (Applied Biosystems). To test the reproducibility of the AFLP technique, two independent DNA extractions were used for all strains. Strain 306 of X. citri pv. citri (Da Silva et al., 2002) was used as a control in each experiment.

Amplified fragments were scored as present (1) or absent (0) to create binary matrices, before analysis using R software (version 2.6.1; R Development Core team). Only bands with an intensity above a preset level (500 relative fluorescence units) were scored. EGD values were calculated from the Dice similarity index and corrected to account for unobserved substitutions by using the standard Jukes–Cantor model, which assumes equal rates of nucleotide substitution between all pairs of bases (Mougel et al., 2002). EGD values were used as distances to construct a weighted neighbour-joining (NJ) tree (Paradis et al., 2004) using R software. The robustness of the tree was assessed by bootstrap (1000 resamplings).
**MLSA.** Four loci, *atpD*, *dnaK*, *efp* and *gyrB* (Ah-You et al., 2009; Boudon et al., 2005), were used for analysis (Table 1). All loci were amplified using the BD advantage 2 polymerase mix kit (Clontech), as recommended by the manufacturer, using primers shown in Supplementary Table S3. Amplifications were performed in a PE9600 thermocycler (Applied Biosystems) as follows: an initial denaturation at 95 °C for 3 min was followed by 35 cycles of denaturation for 1 min at 95 °C, annealing for 1 min at 66 °C and extension for 1 min at 68 °C and a final extension step at 68 °C for 7 min. For a small number of strain–gene combinations, the annealing temperature was modified (63–68 °C) for optimal PCR quality or yield. The amplicons were sequenced by Macrogen or Genome Express (Gibbs et al., 2007), available in MEA 4.1 (Tamura, 1997), implemented in the RDP3 program (Martin 1997), and were evaluated using Kimura’s two-parameter method (Kimura, 1980) with the MEA 4.1 program (Tamura et al., 2007).

Alignments were screened for evidence of recombination, identification of likely parental sequences and localization of possible recombination breakpoints by using a set of seven non-parametric detection programs: RDP (Martin & Rybicki, 2000), Geneconv (Padidam et al., 1999), MaxChi (Maynard-Smith, 1992), Chimera (Posada & Crandall, 2001), bootscanc (Martin et al., 2005a), SiScan (Gibbs et al., 2000) and 3Seq (Boni et al., 2007). The programs are implemented in the RDP3 program (Martin et al., 2005b). The analysis was performed with default settings for the different detection methods and the Bonferroni-corrected P-value cut-off was set at 0.05. Recombination events were accepted when detected with three detection methods or more. Breakpoint positions and recombinant sequences inferred for every detected potential recombination event were checked manually and adjusted where necessary using the extensive phylogenetic and recombination signal analysis features available in RDP3 version beta 15.

Maximum-likelihood (ML) methods were used to infer phylogenetic relationships for each single gene and concatenated genes. The appropriate model of evolution for each dataset was estimated using modeltest 3.8 (Posada, 2006) based on the Akaike information criterion (Posada & Buckley, 2004). The GTR + Γ + I model was selected for *atpD*, *dnaK*, *gyrB* and the concatenated gene dataset, while the TIM + Γ + I model was selected for *efp*. Distance matrices derived from each single-gene and concatenated dataset, taking into account the selected model of evolution, were generated in PAUP* version 4.0b10. Correlations between distance matrices derived from AFLP and MLSA using concatenated sequences were tested pairwise using the Mantel test (Mantel, 1967). All Mantel tests were performed using GenAlEx version 6.1 with 9999 permutations (Peakall & Smouse, 2006).

ML trees, including bootstrap analyses (1000 resamplings), were performed with PhyML software (Guindon & Gascuel, 2003). NJ trees were constructed using PAUP* version 4.0b10 with 1000 replications. Bayesian maximum-likelihood inference was used to reconstruct phylogenies based on the concatenated data using MrBayes version 3.1.1 (Ronquist & Huelsenbeck, 2003). This program implements the Markov chain Monte Carlo (MCMC) algorithm to approximate the posterior probability distribution of a large number of trees. Two runs with four Markov chains were run simultaneously for 5.0 × 10⁶ generations and sampled every 100 generations. Variations in the ML trees were examined graphically using the Tracer software (http://tree.bio.ed.ac.uk/software/tracer/).

The Shimodaira–Hasegawa (S-H) test (Shimodaira & Hasegawa, 1999) was used to determine whether the tree topologies based on each locus fall within the same confidence limits. The significance of differences in the likelihood scores was assessed using a bootstrap test with 1000 replications. This analysis was performed with PAUP* version 4.0b10. The S-H test was also used to compare tree topologies (NJ, ML and Bayesian methods) based on concatenated data.

The molecular clock hypothesis was tested using Tree-puzzle (Schmidt et al., 2002). ML trees were constructed for each or all codon positions using the quartet puzzling algorithm with and without enforcing the molecular clock. Trees were constructed for each gene using optimized parameters determined previously by MODELEST. Likelihood ratio tests (*P* = 0.05) were used to compare trees constructed with and without enforcing the molecular clock. When the molecular clock hypothesis was rejected, the same analysis was performed taking each codon position separately. Mean distances between alleles at synonymous (Dₐ) and non-synonymous (Dₙ) sites were calculated using concatenated data after Jukes–Cantor correction by using MEGA 4.1 (Tamura et al., 2007).

**Table 1.** Results of the S-H test of alternative tree topologies for the genes used for MLSA

<table>
<thead>
<tr>
<th>Tree topology</th>
<th>Likelihood scores (−ln L) for the following datasets</th>
<th>Concatenated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>atpD</em></td>
<td><em>dnaK</em></td>
</tr>
<tr>
<td><em>atpD</em></td>
<td>2479</td>
<td>2667***</td>
</tr>
<tr>
<td><em>dnaK</em></td>
<td>2949***</td>
<td>2439</td>
</tr>
<tr>
<td><em>efp</em></td>
<td>3301***</td>
<td>2676***</td>
</tr>
<tr>
<td><em>gyrB</em></td>
<td>3041***</td>
<td>2641***</td>
</tr>
<tr>
<td>Concatenated–ML</td>
<td>2639</td>
<td>2548***</td>
</tr>
<tr>
<td>Concatenated–NJ</td>
<td>2652</td>
<td>2526***</td>
</tr>
<tr>
<td>Concatenated–MrBayes</td>
<td>2639</td>
<td>2538***</td>
</tr>
</tbody>
</table>

**P<0.01; ***, P<0.001. The lowest (best) likelihood scores are indicated in bold.**
RESULTS

Pathogenicity tests
All pathotype A strains of pv. citri produced erumpent, callus-like lesions a week after inoculation, irrespective of the assayed host species. Pathotype A* and Aw strains produced erumpent, callus-like lesions on Mexican lime, Tahiti lime and alemow a week after inoculation, but not on grapefruit or sweet orange leaves. Very small, blister-like lesions were observed for some strains, sometimes in an irreproducible way, on the latter two citrus species. Detached Mexican lime leaves inoculated with X. citri pv. aurantifolii strains showed erumpent, callus-like lesions similar to those caused by pathotype A, A* and Aw strains. No lesions were recorded on leaves of other Citrus species inoculated with strains of this pathovar. Strains of X. campestris pv. bilvae did not cause canker-like symptoms, but produced extensive water-soaked lesions on Mexican lime leaves. Mean population densities 1 month after inoculation from these lesions ranged from $1 \times 10^7$ to $2 \times 10^7$ c.f.u. per lesion, typical of a compatible interaction. Necrotic lesions were observed on inoculated grapefruit and sweet orange leaves. Such reactions were regarded as incompatible interactions, given that mean population densities recovered from such lesions 1 month after inoculation were $\leq 2 \times 10^5$ c.f.u. per lesion. Differential pathogenicity among strains of X. campestris pv. bilvae was recorded on alemow. Small but consistently developing water-soaked lesions were observed for strain NCPPB 3213 (mean population size $2 \times 10^6$ c.f.u. per lesion) but not for strain NCPPB 1759 (mean population density $5 \times 10^5$ c.f.u. per lesion).

When bacterial multiplication was assayed over time, population densities of three pathotype A* strains (JK2-13, JK143-12 and JS751) were not different from those of pathotype A strains on alemow at 35 days p.i. Population densities recorded from lesions on alemow for strains CFBP 2911 and X2003-01029G at 35 days p.i. were significantly lower (by less than 0.8 log unit) than from other assayed strains. Population densities recorded from 20 to 35 days p.i. on grapefruit for all pathotype A*/Aw strains were significantly lower than those recorded for pathotype A strains. One pathotype A* strain (JS551) from Iran produced small, blister-like lesions on grapefruit. X. citri pv. citri population densities in blister-like lesions recorded 7 to 35 days p.i. were approximately $1 \times 10^5$ c.f.u. per lesion, two log units lower than those recorded from canker-like lesions. Population densities for other pathotype A*/Aaw strains on grapefruit were lower by three to four log units than those recorded for pathotype A strains.

PCR assay
Amplicons of the expected size were obtained with all X. citri pv. citri strains after the second round of PCR. However, no amplification product was visible after the first round of PCR with six pathotype Aw strains from Florida and one pathotype A* strain from India (NCPPB 3608). The nucleotide sequences of the amplicon produced after the second round of PCR for strains X2003-01035G (pathotype Aw) and NCPPB 3608 (pathotype A*) were fully identical (over 233 bp) to that of the target sequence (GenBank accession numbers AE008924.1 and AE008925.1).

AFLP analysis
Cluster analysis, based on 1314 AFLP fragments, placed most Xanthomonas genospecies in distinct lineages, supported by high bootstrap values (>80 %) (Fig. 1). Strains identified as pathovars of X. axonopodis sensu Vauterin et al. (1995) were distributed in six clusters, corresponding to the subgroups 9.1 to 9.6 sensu Rademaker et al. (2005). All subgroups except subgroup 9.6 were supported by bootstrap values above 80 %. Strains of pathovars citri and aurantifolii clustered within X. citri sensu Ah-You et al. (2009) (synonyms X. axonopodis subgroups 9.5 and 9.6), which is consistent with earlier studies. Strains that caused citrus canker on a narrow host range (pathotype A* and Aw strains) were most closely related to X. citri pv. citri strains with a wide host range (pathotype A): strains from the three pathotypes formed a group, which was quite distinct from other X. citri pathovars and supported by a strong bootstrap value (Fig. 1). Within-group EGD values ranged from 0.0010 to 0.0161. Pathotype Aw strains were most closely related to a pathotype A* strain originating from India (NCPPB 3608), with an EGD of 0.0038. Pathotype A* (including Aw strains; n=38) was found to be more diverse than pathotype A (n=73), with median EGD values of 0.00802 and 0.00276, respectively. EGD values between the type strain of X. citri and strains of X. citri pv. aurantifolii ranged from 0.0598 to 0.0693.

When compared to all type strains of Xanthomonas, strains of X. campestris pv. bilvae were most closely related to the X. citri genospecies, with EGD values ranging from 0.0354 to 0.0372. Their relatedness to the X. citri pv. citri strain collection studied here ranged from 0.0322 to 0.0512.

MLSA analysis
Analysed sequence lengths ranged from 389 bp (efp) to 773 bp (gyrB), leading to a total of 2671 bp for the four regions sequenced (Table 2). The number of alleles ranged from 30 (efp) to 42 (gyrB) (Table 2), with a mean of 36.8 alleles per locus. The G+C content for all loci ranged from 63.1 to 64.6 mol%. The nucleotide diversity per site ranged from 0.0228 (efp) to 0.0315 (gyrB). Nucleotide transitions exceeded transversions. All genes showed a mean Ka/Ks <<1, indicating that they were subject to stabilizing selection, conforming to the general requirements of MLSA. Stabilizing selection was further suggested by Tajima’s D values for all genes (Table 2). No recombination event involving strains of X. citri was detected. Recombination was detected for Xanthomonas euvesicatoria for a part of the atpD gene (Supplementary Fig. S1). The detected recombination resulted in a distorted representation of X. axonopodis.
subgroup 9.2 sensu Rademaker et al. (2005), notably increasing the branch length of *X. euvesicatoria* on the tree derived from concatenated sequence data (Fig. 2).

The phylogenetic tree for each locus consistently clustered strains of *X. axonopodis* sensu Vauterin et al. (1995) according to Rademaker’s subgroups (Supplementary Figs S1–S4). However, subgroups 9.1 and 9.4 did not appear as robust groups on trees based on *efp* or *gyrB*. Likelihoods for a given gene computed using the S-H test were statistically different from those computed for the three other loci (Table 1). Most often, likelihoods computed from a single locus were congruent with that derived from concatenated sequence data, with the exception of *dnaK* with all trees inferred from the concatenated datasets. All pathotype A, A* and Aw strains of *X. citri* pv. *citri* as well as *X. campestris* pv. *bilvae* (NCPPB 3213 and NCPPB 1759) consistently clustered within the *X. citri* genospecies, regardless of the gene under analysis. No polymorphic site was found at the four examined loci between pathotype A* and Aw strains. Pathotype A*/Aw strains were most closely related to pathotype A strains. Three sequence types (ST) were identified within *X. citri* pv. *citri*. All pathotype A*/Aw were assigned to ST1, whereas all but one pathotype A strains were assigned to ST2. The two groups could be distinguished on the basis of four synonymous mutations in the *atpD* sequence (one in the first codon position and three in the third codon position), corresponding to a synonymous nucleotide variation of 0.59%. However, the molecular clock hypothesis was rejected for *atpD* (at the third codon position), precluding calculation of age estimates between pathotypes. So far, only strains NCPPB 3562 and LD7-1 (from India and Mali, respectively) were classified as ST3, for which a single synonymous mutation (at the third codon position) was identified in *dnaK*. ST3 differed from ST1 and ST2 by a synonymous nucleotide variation of 0.74 and 0.15%, respectively.

**Fig. 1.** Neighbour-joining tree derived from EGD values (Mougel et al., 2002; Portier et al., 2006) showing the relationships between *X. citri* and other *Xanthomonas* genospecies based on 1314 AFLP markers. Branches with bootstrap values lower than 80% are represented by dotted lines.
Table 2. Sequence variations for four housekeeping genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>gyrB</th>
<th>dnaK</th>
<th>efp</th>
<th>atpD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (bp)</td>
<td>773</td>
<td>762</td>
<td>389</td>
<td>747</td>
</tr>
<tr>
<td>Position in strain 306 (Mb)</td>
<td>0.0</td>
<td>1.8</td>
<td>2.8</td>
<td>4.3</td>
</tr>
<tr>
<td>G + C content (mol%)</td>
<td>63.1</td>
<td>63.5</td>
<td>64.6</td>
<td>64.6</td>
</tr>
<tr>
<td>Ts/Tv*</td>
<td>3.532 (0.857)</td>
<td>2.893 (0.999)</td>
<td>1.256 (0.512)</td>
<td>1.715 (0.636)</td>
</tr>
<tr>
<td>Mean nucleotide identity (minimum)</td>
<td>0.968</td>
<td>0.969</td>
<td>0.977</td>
<td>0.976</td>
</tr>
<tr>
<td>Ka†</td>
<td>0.0014 (0.0028)</td>
<td>0.0016 (0.0015)</td>
<td>0.0037 (0.0046)</td>
<td>0.0054 (0.0051)</td>
</tr>
<tr>
<td>Ks†</td>
<td>0.1418 (0.0899)</td>
<td>0.1389 (0.0740)</td>
<td>0.0989 (0.0884)</td>
<td>0.0992 (0.0589)</td>
</tr>
<tr>
<td>Kd/Ks†</td>
<td>0.0106 (0.0103)</td>
<td>0.0115 (0.0397)</td>
<td>0.0374 (0.0598)</td>
<td>0.0544 (0.0428)</td>
</tr>
<tr>
<td>Alleles (n)</td>
<td>42</td>
<td>39</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>Variable sites (sites leading to synonymous mutations) (n)</td>
<td>146</td>
<td>115</td>
<td>60</td>
<td>129</td>
</tr>
<tr>
<td>Nucleotide diversity per site</td>
<td>0.0315</td>
<td>0.0313</td>
<td>0.0228</td>
<td>0.0271</td>
</tr>
<tr>
<td>Tajima’s D‡</td>
<td>−1.1512</td>
<td>−0.5289</td>
<td>−1.4114</td>
<td>−1.1874</td>
</tr>
</tbody>
</table>

*D*Transition/transversion ratio determined using Kimura’s two-parameter method (Kimura, 1980). Values in parentheses are standard deviations. †Synonymous and non-synonymous substitution rates determined using the method of Nei & Gojobori (1986). Values in parentheses are standard deviations. ‡Not significant (P>0.10).

Phylogenetic trees were inferred from concatenated sequences using the NJ, ML and Bayesian methods based on evolution model GTR+Γ+I (Fig. 2 and Supplementary Figs S5 and S6). The average branch length produced from the Bayesian analysis after estimating the lambda parameter from the ML tree (lambda=293) was similar to those calculated for the NJ and ML trees, at 0.0031, 0.0015 and 0.0023, respectively. All methods produced congruent trees (Table 1), as indicated using S-H tests of alternative tree topologies. As shown with the ML tree (Fig. 2), major branches were supported by high bootstrap values (>80%) and strain grouping matched the X. axonopodis subgroups proposed by Rademaker et al. (2005). Pathotype A, A* and Aw strains causing citrus canker always grouped within X. citri and were highly related, irrespective of the phylogenetic inference method employed. All strains of X. citri pv. citri were most distantly related to X. citri pv. aurantifolii, which was consistent with earlier studies (Cubero & Graham, 2002; Rademaker et al., 2005). Pathotype A* and Aw strains shared identical sequence types and were distinguishable from pathotype A strains. X. campestris pv. bilvae strains clustered within the X. citri genospecies. They were most closely related to X. citri pv. glycines (LMG 712).

Distances between pathotype A and pathotype A*/Aw strains calculated with the GTR+Γ+I model, based on concatenated sequences, ranged from 0.0015 to 0.0019, while genetic distances from other pathovars in the same genospecies ranged from 0.0015 to 0.0235 and were greater than 0.0273 with subgroups 9.1 to 9.4 of X. axonopodis. The genetic distance between the type strain of X. citri and strains of X. campestris pv. bilvae was 0.0030. The genetic distance between the type strain of X. citri and strains of X. citri pv. aurantifolii ranged from 0.0211 to 0.0235. Correlation values (w=0.861, P<0.001) obtained from the Mantel test suggested that data derived from AFLP and MLSA using concatenated sequences were highly congruent.

DISCUSSION

The pathovar classification of bacterial plant pathogens (Dye et al., 1980; Young et al., 1992) is an infrasubspecific classification, applied by reference to host range or capacity to cause distinctive symptoms. Pathogenicity tests are, therefore, an essential part of this classification. However, an increasing number of contributions have demonstrated that classifications based on genetic fingerprinting or sequencing methods, which correlate well with DNA–DNA hybridization data (e.g. AFLP, rep-PCR and/or MLSA), also often provide an adequate reflection of the pathovar classification in the genus Xanthomonas (Ah-You et al., 2007, 2009; Louws et al., 1999; Rademaker et al., 2000, 2005, 2006; Scortichini & Rossi, 2003). This indicates that pathovars are distinct genetic entities, even at the genome level. This differentiation can be explained by the separation of ecological niches induced by host specialization, which has led to different selective pressures and to de facto environmental isolation. It also suggests that the host specialization processes outlined at the pathovar level occurred a long time ago. However, the date of such separation in the course of evolution has not been documented.

Host specialization associated with subtle genetic polymorphism has sometimes been referred to as pathotypes, an infrapathovar classification (Civerolo & Fan, 1982; Verdier et al., 1998). Our study aimed primarily at characterizing phylogenetic relationships between pathotypes of X. citri pv. citri and the genetically related pathovar...
To achieve this, we used AFLP and MLSA, techniques that have been given emphasis for species delineation in bacterial taxonomy (Stackebrandt et al., 2002).

In this study, we show that bacterial strains that were isolated from members of two rutaceous host genera (Aegle and Feronia) in India, currently classified as X. campestris, are closely related to the type strain of X. citri but distantly
related to the type strain of *X. campestris*. Based on AFLP and MLSA, these strains were more closely related to *X. citri* pv. *citri* than to *X. citri* pv. *aurantifolii*. Genetic distances between *X. citri* and *X. campestris* pv. *bilvae* derived from these two techniques corresponded to ΔTm values <2.0 °C (Ah-You et al., 2009). Therefore, these strains should be reclassified within the *X. citri* genospecies as *X. citri* pv. *bilvae*. Pathogenicity tests showed that strains of *X. citri* pv. *bilvae* produced extensive lesions on attached Mexican lime leaves and multiplied markedly within the mesophyll, therefore confirming an earlier report from India (Patel et al., 1953). The fact that the pathovars *citri* and *bilvae* are commonly pathogenic to *Aegle marmelos*, *Feronia limonia* and *Citrus aurantifolia*, but have clearly distinct symptomatologies, justifies their classification as two different pathovars of *X. citri* (Dye et al., 1980). This finding indicates that xanthomonads that produce bacterial spot-like lesions on rutaceous host species are not restricted to *X. axonopodis* subgroup 9.2 (synonym *Xanthomonas alfalfae*), as previously thought (Rademaker et al., 2005; Schaad et al., 2005).

Interestingly, genetic distances derived from AFLP and MLSA were highly correlated. These techniques show that all strains that cause canker-like lesions on citrus, classified as pathovars *aurantifolii* and *citri* (Gabriel et al., 1989), are members of the *X. citri* genospecies *sensu* Ah-You et al. (2009). All three pathotypes described within pathovar *citri* (Sun et al., 2004; Vernière et al., 1998) were closely related genetically, confirming previous DNA–DNA hybridizations performed on a limited number of strains (Sun et al., 2004). Whatever the technique used, all strains of *X. citri* pv. *citri* with a narrow host range (pathotypes A* and A*) were more closely related to strains with a wide host range (pathotype A) than to any other pathovar within *X. citri*. MLSA showed that all strains belonging to pathotypes A* or A” were identical, suggesting that they shared a fairly recent common ancestor. These strains could also not be distinguished on the basis of pathogenicity tests. They shared a similar host range, restricted to Mexican lime, Tahiti lime and alemow, with no differences in symptomatology. Some pathotype A* strains caused small, blister-like reactions on sweet orange and grapefruit, but these reactions were not observed reproducibly for some strains, making this criterion impossible to use for pathotype classification. The weaker *in planta* multiplication of strains producing blister-like reactions suggests that such reactions should not be considered typical compatible interactions, consistent with earlier reports (Ah-You et al., 2007; Egel et al., 1991). Our results on the host range of pathotype A* and A” strains disagree with those reported by Sun et al. (2004). These authors similarly scored disease reactions caused by pathotype A* strains on Mexican lime and sweet orange. Sweet orange was described as the differential host species for distinguishing pathotypes A* and A” (Sun et al., 2004). In the present study, we show that sweet orange leaves inoculated with pathotype A* strains show no or very limited blister-like reactions but no canker-like reactions. Consequently, our data showed no clear difference in the host ranges of pathotype A* and A” strains, consistent with the results of Brunings & Gabriel (2003).

Furthermore, one feature of pathotype A” strains that was considered to be unique was their lack of amplicon production from primers specific to *X. citri* pv. *citri* (Sun et al., 2004). In this study, we showed that this feature was shared with some pathotype A* strains originating from India. All the strains for which no amplicon was visible using these primers did produce a visible amplicon with internal primers. The DNA sequence of a 233 bp amplicon portion obtained from a pathotype A” strain was fully identical to that of the target fragment, suggesting that the lack of PCR amplification with external primers may be due to polymorphisms at the binding sites of external primers. We conclude that pathotype A” is a junior synonym of pathotype A*, a group of strains with a wide genetic diversity, and that the latter designation should be preferred. An analysis of the genetic diversity of 234 strains of *X. citri* pv. *citri* originating from Asia, the likely centre of origin of the pathogen, by AFLP and insertion sequence ligation-mediated PCR targeting three insertion sequences suggested that pathotype A* (including A” strains) is a group of strains that shows a wider genetic diversity than pathotype A, suggesting a longer evolutionary history of pathotype A* strains (Bui Thi Ngoc et al., 2009).

**ACKNOWLEDGEMENTS**

We thank J. Young and E. L. Civerolo for helpful discussion and C. Boyer, K. Vital and V. Ledoux for their technical assistance. The European Union (FEOGA and FEDER), Conseil Régional de La Réunion and CIRAD provided financial support.

**REFERENCES**


