Applied and Environmental Microbiology, June 2008, p. 3652–3657 Vol. 74, No. 12
0099-2240/08/$08.00+0 doi:10.1128/AEM.02386-07
Copyright © 2008, American Society for Microbiology. All Rights Reserved.

# Hypervariations of a Protease-Encoding Gene, PD0218 (*pspB*), in *Xylella fastidiosa* Strains Causing Almond Leaf Scorch and Pierce's Disease in California[▽]

J. Chen,[1]* E. Civerolo,[1] K. Tubajika,[2] S. Livingston,[3] and B. Higbee[4]

*USDA-ARS, Parlier, San Joaquin Valley Agricultural Sciences Center, 9611 S. Riverbend Avenue, Parlier, California 93648[1]; USDA-APHIS, Center for Plant Health Science and Technology, Pest Epidemiology and Risk Assessment Laboratory, 1730 Varsity Drive, Suite 300, Raleigh, North Carolina 27606[2]; Department of Plant Pathology, University of California, Davis, California 95616[3]; and Paramount Farming Co., 33141 E. Lerdo Highway, Bakersfield, California 93308[4]*

Received 23 October 2007/Accepted 10 April 2008

***Xylella fastidiosa* is a gram-negative plant pathogenic bacterium that causes almond leaf scorch disease (ALSD) and Pierce's disease (PD) of grape in many regions of North America and Mexico. Of the two 16S rRNA gene genotypes described in California, A genotype strains cause ALSD only and G genotype strains cause both PD and ALSD. While G genotype strains cause two different diseases, little is known about their genetic variation. In this study, we identified a putative protease locus, PD0218 (*pspB*), in the genome of *X. fastidiosa* and evaluated the variation at this locus in *X. fastidiosa* populations. PD0218 contains tandem repeats of ACDCCA, translated to threonine and proline (TP), upstream of the putative protease conserved domain. Among 116 *X. fastidiosa* ALSD and PD strains isolated from seven locations in California, tandem repeat numbers (TRNs) varied from 9 to 47, with a total of 30 TRN genotypes, indicating that *X. fastidiosa* possesses an active mechanism for contracting and expanding tandem repeats at this locus. Significant TRN variation was found among PD strains (mean = 29.9), which could be further divided into two TRN groups: PD-$G_{small}$ (mean = 17.3) and PD-$G_{large}$ (mean = 44.3). Less variation was found in ALSD strains (mean = 21.7). The variation was even smaller after ALSD strains were subdivided into the A and G genotypes (mean = 13.3, for the G genotype; mean = 27.1, for the A genotype). Genetic variation at the PD0218 locus is potentially useful for sensitive discrimination of *X. fastidiosa* strains. However, TRN stability, variation range, and correlation to phenotypes should be evaluated in epidemiological applications such as pathotype identification and delineation of pathogen origin.**

*Xylella fastidiosa* is a gram-negative and nutritionally fastidious plant pathogenic bacterium that causes almond leaf scorch disease (ALSD) and Pierce's disease (PD) of grapevine (*Vitis vinifera*). The two diseases have recently reemerged as threats to the production of both almonds and grapes in California. ALSD and PD strains of *X. fastidiosa* were previously reported to be pathologically similar (8, 20). Recent studies showed that *X. fastidiosa* strains associated with ALSD in California are composed of two genetically and pathologically distinct groups (1, 5, 14). Based on single-nucleotide polymorphisms (SNPs) in the 16S rRNA gene sequences, Chen et al. (5) assigned the strains causing only ALSD to the A genotype and strains causing both ALSD and PD to the G genotype. While G genotype strains cause two different diseases, little is known about their genetic variation.

During host-pathogen interactions, plant pathogens produce a number of extracellular enzymes, including proteases, which have been extensively studied in host-pathogen interactions (7, 12, 16, 22, 32). Proteases participate in a wide range of cellular functions, including pathogenicity, and have been detected in *X. fastidiosa* strains (11, 31). Recently, a group of serine proteases were characterized from in vitro cultures of citrus variegated chlorosis disease and PD strains of *X. fastidiosa* (19). A cysteine protease was found to be differentially expressed in a nonpathogenic *X. fastidiosa* strain (21).

The PD0218 (*pspB*) locus of *X. fastidiosa* Temecula-1 encodes a putative serine protease that contains tandem repeats of ACDCCA, translated to threonine and proline (TP), upstream of the protease conserved domain (29). Orthologs of PD0218 containing variable tandem repeat numbers (TRNs) also were found in the genomes of *X. fastidiosa* strains Ann-1, Dixon, and 9a5c (2, 24, 29). While it has not been established for *X. fastidiosa*, it has been documented that other bacteria can change TRNs to modify the function of the expressed proteins during environmental adaptation, new host colonization, and pathogenicity (10, 15, 17, 18, 25, 27, 28).

In this study, a total of 116 *X. fastidiosa* strains were collected from almond orchards and grape vineyards in seven different locations in California. We identified the TRN variation at the PD0218 locus of *X. fastidiosa* populations and evaluated the TRN variation within and between the bacterial hosts and geographical locations. The implications of TRN variation for the pathogen's biology and disease epidemiology are discussed.

## MATERIALS AND METHODS

**Bacterial strains.** *X. fastidiosa* strains were isolated from ALSD-affected almond trees and PD-affected grapevines in California at the Coachella Valley

* Corresponding author. Mailing address: USDA-ARS, Parlier, San Joaquin Valley Agricultural Sciences Center, 9611 S. Riverbend Ave., Parlier, CA 93648. Phone: (559) 596-2924. Fax: (559) 596-2921. E-mail: Jianchi.chen@ars.usda.gov.

▽ Published ahead of print on 2 May 2008.

```
A
NP_778460(PD0218) MKNMKTAFFPGSILVLTLVAFLSACGGSGLVRPPtptptptptptptptptptptptptp 60
NP_778548(PD0313) METMKTIFAARTILASALAAALSACGDP-LVRSEtptpttpttpt-tpttpttpttpttp 58
NP_779165(PD0950) MQKIKNKFIVRTILATTVTTVLSACGGG-LVRIDtptptptptptptptptptptptptp 59
                  *:.:*. *   :**. ::.: *****.  ***  *****...*** ***...***...**

NP_778460(PD0218) tptptptptptptptptp---------------------------------------LPV 81
NP_778548(PD0313) t-tptp---------------------------------------------------LPA 66
NP_779165(PD0950) tptptptptptptptptptptptptptptptptptptptptptptptptptptptpTLPP 119
                  * ****                                                   **

NP_778460(PD0218) LDPALAAKFRVHLDVTNTEPAHTLGFTGKGYMVGVVDSGVNRNHVTLHDQVIHN---NVY 138
NP_778548(PD0313) LDPALAAEFRVHLNGTNTEPAHTLGFTGKGYMIGVVDSGVNRNHVTLHDQVIHN---NVY 123
NP_779165(PD0950) LQPSLAAKIGRHLELTNTEAAHALGLTGSGFIIGMVDTGVNRNNIALQGKVIRNGVYNFY 179
                  *:*:***::   **: ****.**:**:**.*:::*:**:*****::*:.:**:*   *.*

B
PD0218   ATGAAAAATATGAAAACGGCATTTTTTCCCGGTAGCATCCTGGTATTAACGTTAGTTGCT 60
PD0313   ATGGAAACGATGAAAACGATATTTGCTGCCAGGACGATCTTGGCATCCGCATTGGCTGCT 60
PD0950   ATGCAAAAGATTAAAAATAAATTTATTGTTAGAACGATCTTGGCAACAACCGTAACTACT 60
         *** ***  ** ****    ****  *    * *  *** *** *    *  *   * **

              pspB-256f
          5'- GAGTGCCTGTGGTGGTA - 3'
PD0218   TTCTTGAGTGCCTGCGGTGGTAGCGGGTTAGTTCGTCCGCCAacgccaacgccaacgcca 120
PD0313   GCTTTGAGTGCCTGCGGCGATC---CATTAGTTCGCTCAGAGacaccgacaccgacgaca 117
PD0950   GTCTTGAGTGCCTGTGGTGGTC---GATTAGTTCGTATAGATacacctacaccgacaccc 117
           *********** ** * *     ********        ** ** ** ** **  *

PD0218   acaccaacaccaacaccaacaccaacaccaacaccaacaccaacaccaacgccaacgcca 180
PD0313   ccgacgacaccgacg---------------------------acaccgacgacaccgacg 150
PD0950   acacctacacccacaccgacacccacaccgaca------cccacaccgacacccacaccg 171
          *  * ***** **                            ***** **  *  *  *

PD0218   acaccaacaccaacaccaacaccaacaccaacaccaacaccaacaccaacacca------ 234
PD0313   acaccgacg---------------acaccgacgacaccgacgacaccgacacca------ 189
PD0950   acacccacacctacacccacacctacacccacacccacacccacacccacacctacaccc 231
         ***** **                ***** **  *  *  * ***** *****

PD0218   ------------------------------------------------------------ 234
PD0313   ------------------------------------------------------------ 189
PD0950   acacctacacccacacctacacccacacctacacccacacctacacccacacctacaccc 291

PD0218   ---------------------------------------------------------TTA 237
PD0313   ---------------------------------------------------------TTG 192
PD0950   acacctacacccacacccacacccacacctacacccacacctacacctacacccACATTG 351
                                                                   **

                             pspB-256r
                       5'- CGAAACTTGGCAGCTAACG -3'
                       3'- CGTTAGCTGCCAAGTTTCG -5'
PD0218   CCTGTTTTGGATCCGGCGTTAGCTGCCAAGTTTCGCGTGCAC 279
PD0313   CCTGCTTTGGATCCGGCGTTAGCTGCCGAGTTTCGGGTGCAC 234
PD0950   CCCCCACTGCAACCATCATTAGCTGCGAAGATCGGGCGGCAT 393
         **     ** * **  * ********  ** *  *   ***

C
                                   pspB-256f                                                                         pspB-256r
                              5'- GAGTGCCTGTGGTGGTA -3'                                                  5'- CGAAACTTGGCAGCTAACG -3'
                                                                                                         3'- CGTTAGCTGCCAAGTTTCG -5'
NZ_AAAL02000012(Dixon)    TTGAGTGCCTGTGGTGGTAGCGGGTTAGTTCGTCCGCCA(ACDCCA)30TTGCCTGATTTGGATCCGGCGTTAGCTGCCAAGTTTCGCGTGCAC
PD0218(Temecula-1)        TTGAGTGCCTGCGGTGGTAGCGGGTTAGTTCGTCCGCCA(ACDCCA)22TTACCTGTTTTGGATCCGGCGTTAGCTGCCAAGTTTCGCGTGCAC
NZ_AAAM03000002(Ann-1)    TTGAGTGCCTGTGGTGGTAGCGGGTTAGTTCGTCCGCCA(ACDCCA)12TTGCCTGCTTTGGATCCGGCGTTAGCTGCCAAGTTTCGCGTGCAC
XF0267(9a5c)              TTGAGTGCCTGTGGTGGTAGTGGGTTAGTTCGCCCATCG(ACACCA)29TTGCCTGCTTTGGATCCGGCGTTAGCTGCCAAGTTTCGGGTGCAC
                            ********* *******                                               *******************
```

FIG. 1. Alignments of amino acid sequences (A) translated from nucleotide sequences (B) of PD0218, PD0313, and PD0950 from *X. fastidiosa* strain Temecula-1 in the vicinity of tandem repeat regions. Tandem repeat regions are in lowercase letters. (C) Nucleotide sequence alignment of PD0218 and its closest homologs from *X. fastidiosa* strains Ann-1, Dixon, and 9a5c in the region flanked by primers pspB-256f and pspB-256r. In the last line of each comparison, stars indicate amino acid or nucleotide identity, single and double dots indicate amino acid similarity, and empty spaces represent dissimilarity. PCR primer sequences and the corresponding regions are in bold type.

(Riverside County), Fresno County, Kern County, the Napa Valley (Napa County), Santa Maria (Santa Barbara County), and the Temecula Valley (Riverside County) between August and November during 2003 to 2005. At each location, symptomatic leaves were collected from randomly selected orchards or vineyards within an approximately 20-km radius. One set of samples was systematically collected from a 14.6-ha almond orchard in Kern County. Isolation procedures were described previously (5). All bacterial strains were cultured on solid PW medium (9) at 28°C for 10 to 14 days. Bacterial cells were suspended in sterile water and used as DNA templates for PCR. For long-term storage, bacterial strains were grown in PW broth and stored at −82°C with the addition of 15% glycerol. *X. fastidiosa* strains Temecula-1 (causing PD) and Dixon (causing ALSD) were included as controls. Both strains were collected from California and provided by A. H. Purcell (University of California, Berkeley). Strain Temecula-1 was isolated from a grapevine in the Temecula Valley in Riverside County, and strain Dixon was from an almond tree in Solano County.

**Primers and PCR procedure.** The sequence of PD0218 (*pspB*) in Temecula-1 (AE009442) was obtained from the GenBank database at the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov). PCR primers flanking the tandem repeat region of PD0218 were designed with Primer 3 software (23). The desired midpoint temperature was set to 60 ± 3°C. The designed primer set, pspB-256f/pspB-256r (5′-TGAGTGCCTGCGGTGGTA-3′/5′-CGAAACTTGGCAGCTAACG-3′), was used for DNA amplification as described previously (5). Briefly, the PCR (25 μl) was carried out with the TaKaRa *Taq* (Hot Start Version) kit (Takara Bio Inc., Shiga, Japan). The reaction mixture contained 2.5 μl of 10× DNA polymerase buffer, 2.5 μl of deoxynucleoside triphosphates (2.5 mM each deoxynucleoside triphosphate), 0.5 μl of each of the 10 μM forward and reverse primers, 1 μl of bacterial suspension (~$10^3$ CFU), 0.2 μl of *Taq* DNA polymerase (5 U/μl), and 18.3 μl of $H_2O$. Amplification was conducted in an MJ Research DNA Engine Tetrad 2 with initial denaturation at 96°C for 10 min, followed by 30 cycles of denaturation at 96°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 30 s. Amplification products were then stored at 4°C. Amplicons were resolved by 1.5% agarose gel electrophoresis and visualized by ethidium bromide staining under UV light. For 16S rRNA gene SNP genotype determination, the multiplex PCR procedure (5) was followed. The PCR procedure was the same as that described above, except that primers Teme150fc (5′-TCTACCTTATCGTGGGGGAC-3′) and Teme454rg (5′-AACAACTAGGTATTAACCAATTGCC-3′), specific to the G genotype, and primers Dixon454fa (5′-CCTTTTGTTGGGGAAGAAAA-3′) and Dixon1261rg (5′-TAGCTCACCCTCGCGAGATC-3′), specific to the A genotype, were used. A genotype strains were identified by the presence of an ~800-bp DNA fragment, and G genotype strains were identified by the presence of an ~300-bp DNA fragment.

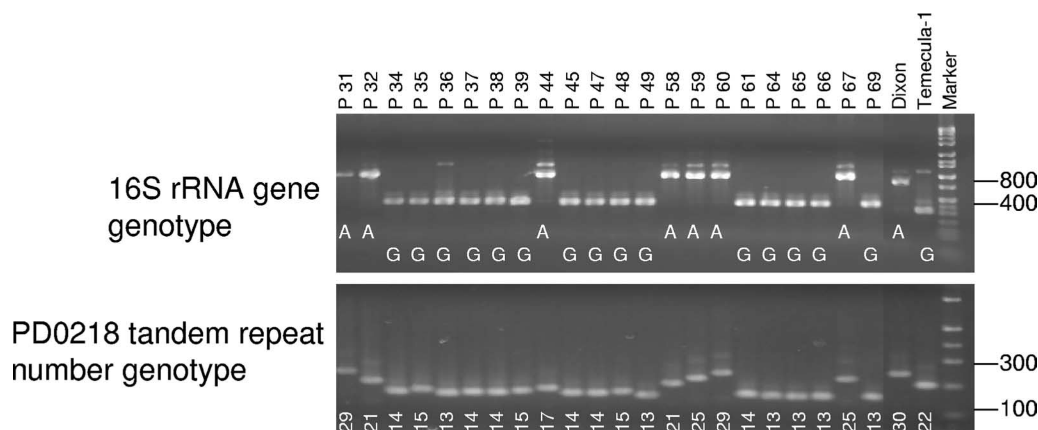


FIG. 2. Results of PCR analyses of *X. fastidiosa* strains isolated from ALSD-affected trees in an orchard in Kern County, CA. Primer sets Teme150fc/Teme454rg and Dixon454fa/Dixon1261rg were used to determine 16S rRNA gene genotypes. Primer set pspB-256f/pspB-256r was used to reveal PD0218 TRN genotypes. The values on the right are sizes of molecular markers in base pairs. Identified at the top are the strains and the DNA size marker. Letters at the bottom of the upper panel indicate 16S rRNA gene genotypes. The values at the bottom of the lower panel are TRNs at the PD0218 locus.

**DNA sequencing and analysis.** PCR amplicons were sequenced directly in both orientations in a 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA) with the same primers used for PCR. Homologs of PD0218 were also downloaded from the genome sequences of *X. fastidiosa* strains Temecula-1, Dixon (AAAL02000000), Ann-1 (AAAM03000000), and 9a5c (AE003849) in the GenBank database. Tandem repeats were counted manually. Homologous genes were identified through BLASTn and BLASTx searches against the GenBank database with PD0218 as a query. Multiple-sequence alignments were performed with the "Query-anchored with identities" option within the NCBI BLAST tool and with the CLUSTAL W program (26) through the web service of the European Bioinformatics Institute (http://www.ebi.ac.uk).

## RESULTS

Annotation of the genome sequence of *X. fastidiosa* Temecula-1 identified three putative serine protease genes: PD0218 (2,967 bp), PD0313 (2,907 bp), and PD0950 (3,078 bp) (29). The sequence of PD0218 was divided into four regions coding for a signal peptide, a string of tandem repeats, a conserved serine-protease domain, and an outer membrane protein domain with an autotransporter related to the type V secretion pathway. Figure 1A shows the amino acid repeats of $[TP]_{22}$ encoded by PD0218, $PT(PTT)_8PT$ encoded by PD0313, and $[TP]_{41}$ encoded by PD0950. For PD0218, the corresponding nucleotide repeat is $[AC(A/G)CCA]_{22}$ (Fig. 1B). The amino acid sequence encoded by PD0218 is 90% identical to that encoded by PD0313 and 70% identical to that encoded by PD0950. Most of the variation is in the TP repeat regions due to different TRNs (Fig. 1A). Similarly, major variations at the nucleotide level are from the tandem repeat region (Fig. 1B).

*X. fastidiosa* strains Dixon and 9a5c have three PD0218 homologs, and strain Ann-1 has four. In all cases, the encoded amino acid sequence identities were greater than 70% and most of the variation was due to difference in TRNs (data not shown). Amino acid sequences encoded by the 13 open reading frames were grouped into three clusters. Each cluster contained one sequence from each of the four *X. fastidiosa* genomes, with the exception of strain Ann-1, where two open reading frames (ZP_00684432 and ZP_00683495) were in the same cluster (data not shown).

Primer set pspB-256f/pspB-256r generated a single predominant product from all of the *X. fastidiosa* strains tested, suggesting that the primer set was specific to a single locus (Fig. 2). A comparison of pspB-256f and pspB-256r in PD0218 to the corresponding sequences in PD0313 and PD0950 is shown in Fig. 1B. Significant SNPs at both pspB-256f and pspB-256r prevented efficient amplification, if any, from PD0313 and PD0950. Comparison of sequences flanking the tandem repeat region of PD0218 homologs from the four bacterial genomes is presented in Fig. 1C. At the corresponding primer location, strain Temecula-1 was different from strains Ann-1, Dixon, and 9a5c by one SNP in the middle of primer pspB-256f. This SNP did not affect DNA amplification from strain Dixon and other A genotype strains (Fig. 2).

Among the 116 *X. fastidiosa* ALSD and PD strains collected from seven locations in California, TRNs varied from 9 to 47 (Table 1), with a total of 30 TRN genotypes (Fig. 3). TRN means, standard deviations, and ranges are summarized in Table 1. By combining host origin and 16S rRNA gene genotype information with TRN data, four TRN groups were de-

TABLE 1. Summary of means, standard deviations, and ranges of TRNs at the PD0218 locus of *X. fastidiosa* ALSD and PD strains from California

| Host(s), group(s) | 16S rRNA gene SNP genotype | No. of strains | Mean TRN | SD | TRN range |
|---|---|---|---|---|---|
| Grape | | | | | |
| PD-$G_{small}$[a] | G | 24 | 17.3 | 4.1 | 12–23 |
| PD-$G_{large}$[b] | G | 21 | 44.3 | 1.2 | 43–47 |
| PD | G | 45 | 29.9 | 14.0 | 12–47 |
| Almond | | | | | |
| ALSD-G | G | 28 | 13.3 | 2.0 | 9–16 |
| ALSD-A | A | 43 | 27.1 | 3.8 | 14–33 |
| ALSD | G + A | 71 | 21.7 | 7.5 | 9–33 |
| Grape + almond, PD + ALSD | G + A | 116 | 24.9 | 11.2 | 9–47 |

[a] Referred to the smaller TRN range of 12 to 23.
[b] Referred to the larger TRN range of 43 to 47.

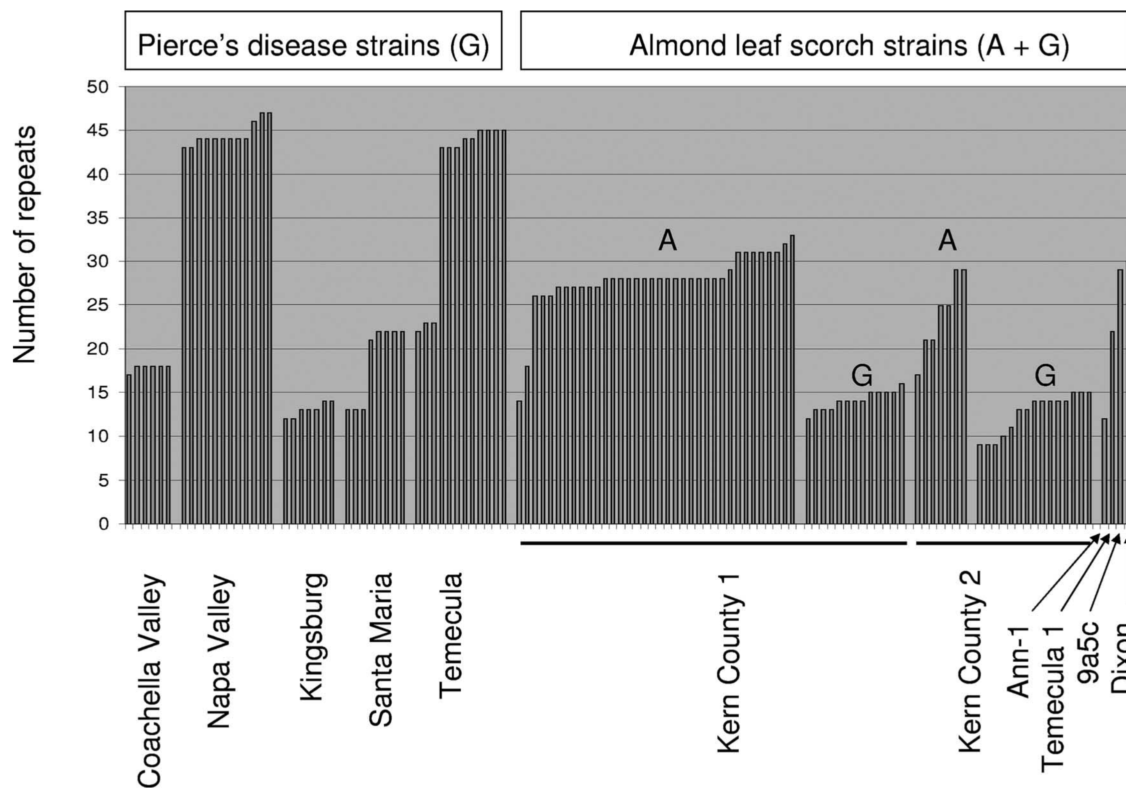


FIG. 3. Distribution of TRNs from *X. fastidiosa* ALSD and grape PD strains isolated from different geographic locations in California. TRNs of strains Ann-1, Temecula-1, 9a5c, and Dixon were obtained from the corresponding whole-genome sequences available in the GenBank database. Statistics for each location: Coachella Valley, mean = 17.8 ± 0.4, *n* = 6; Napa Valley, mean = 44.5 ± 1.4, *n* = 12; Kingsburg, mean = 13.0 ± 0.8, *n* = 7; Santa Maria, mean = 18.5 ± 4.6, *n* = 8; Temecula, mean = 38.8 ± 9.7, *n* = 12; Kern County 1A, mean = 27.8 ± 3.4, *n* = 36; Kern County 1G, mean = 14.1 ± 1.1, *n* = 13; Kern County 2A, mean = 23.9 ± 4.5, *n* = 7; Kern County 2G, mean = 12.6 ± 2.3, *n* = 15.

lineated: PD-$G_{small}$, PD-$G_{large}$, ALSD-A, and ALSD-G (Table 1). Strains in groups PD-$G_{small}$ and ALSD-G had small TRNs (<25) and the same 16S rRNA gene genotype and were pathologically similar (5). Therefore, the two groups could be merged into a single group, namely, the PD-$G_{small}$–ALSD-G group. Five ALSD-A strains had TRNs smaller than 25, falling within the range of the PD-$G_{small}$–ALSD-G group (Fig. 3). TRNs from Temecula-1 and Dixon control sequences were 22 and 30, respectively, identical to those deposited in GenBank.

TRNs varied among *X. fastidiosa* strains from each geographical location (Fig. 3). The lowest variation was in strains from the Coachella Valley (range = 18 to 17; mean = 17.8 ± 0.4; *n* = 6), and the largest variation was in strains from Temecula (range = 46 to 23; mean = 38.8 ± 9.7; *n* = 12). Within the ALSD strains from the same almond orchard (Kern County 1), most of the A genotype strains (34/36) were in group ALSD-A. One genotype (TRN = 28) was predominant (16/36 = 44% among A genotype strains or 16/49 = 33% among both A and G genotype strains).

## DISCUSSION

The high degree of sequence similarity among PD0218, PD0313, and PD0950, as well as their homologs, indicates that all of these putative protease genes likely have the same ancestral origin. The presence of three phylogenetic clusters, each containing at least one gene from *X. fastidiosa* strains Temecula-1, Dixon, Ann-1, and 9a5c, suggests that, during evolution, the ancestral gene was triplicated before host specialization. While the true function of PD0218 needs to be validated experimentally, sequence data indicate that it encodes a protease that is possibly exported extracellularly through a type V secretion pathway. PD0218 may be involved in host-pathogen interactions and therefore deserves further mutation and pathogenicity analyses.

TRNs at the PD0218 locus could be described as hypervariable (Fig. 3), as supported by the total of 30 TRN genotypes from the 116 *X. fastidiosa* strains. It is possible that more TRN genotypes could be found if more strains were evaluated. In addition, multiple TRN genotypes were found at all seven geographic locations studied. These findings indicate that *X. fastidiosa* possesses an active mechanism of contracting and expanding tandem repeats in PD0218. Furthermore, we observed that at least two TRN genotypes were present simultaneously in *X. fastidiosa* strains infecting the same almond tree (data not shown). It was also reported that five TRN genotypes of *X. fastidiosa* citrus variegated chlorosis disease strain were isolated from the same citrus tree (6).

The presence of a predominant TRN genotype (Fig. 3, Kern County 1) suggests that some selection pressure was acting on this bacterial population. Possible selection pressures include host resistance, host physiological status, nutrient composition, and endophytic microbial community. However, we did not

determine if TRN predominant genotypes would change temporally (during a growing season). Data presented in Fig. 3 show that TRN genotypes varied spatially (from location to location).

The molecular mechanism of TRN variation in *X. fastidiosa* is unknown. In general, DNA strand slippage is believed to invoke tandem repeat variations (3, 27, 30). Important issues for future studies include when and where (i.e., other plant hosts, insect vectors, etc.) TRN variations occur, as well as if TRNs vary with passage in culture or in one strain grown on different media, such as solid versus liquid media. In the case of PD0218, the tandem repeats are part of the protease gene. Although it has not been tested experimentally, it is speculated that a TRN change will lead to a change in gene length, resulting in modification of the expressed protein and possibly alteration of its stability, substrate specificity, and enzyme activity.

The ability to adapt to environmental changes is critical for bacterial population survival and expansion (4). One example is the alpha C protein, a surface-associated antigen from human pathogenic group B streptococci (18). The gene (*bca*) for this protein contains a series of tandem repeats that express a protective epitope. Strains from human maternal and neonatal pairs differed in the number of repeats, resulting in differences in host immunity evasion (13, 15). Our data on G genotype strains show that grapes had both PD-$G_{small}$ and PD-$G_{large}$ strains and almonds had only ALSD-G–PD-$G_{small}$ strains. We analyzed the G genotype strains from two additional almond orchards in Fresno County, and only strains in the ALSD-G–PD-$G_{small}$ group were found (data not shown). Whether this is a result of host adaptation of G genotype *X. fastidiosa* remains to be studied. Nevertheless, this is the first report on genetic variation within PD strains and between PD and ALSD strains causing both ALSD and PD (G genotype).

The hypervariation of the PD0218 locus showed its potential for use in the sensitive discrimination of *X. fastidiosa* strains. However, TRN stability, variation range, and occurrence in various plant hosts need further study. For example, model strain Temecula-1 has a TRN of 22 at the PD0218 locus. The frequency of this TRN genotype in Temecula was 1/12 or 0.08, very similar to the frequency of 0.09 (4/45) for all 45 PD strains together. Nine (75%) of 12 PD strains from Temecula had TRNs larger than 40 (Fig. 3), suggesting that the Temecula-1 genome sequence may not be representative of the majority of PD strains in Temecula or California at this locus. Another issue is that if TRN genotypes were generated within an orchard or vineyard and stabilized in a specific environment, it may not be appropriate to delineate pathogen origins among different locations based on TRN similarity.

## ACKNOWLEDGMENTS

Part of this research was supported by a UC-DANR PD research grant.

We thank R. Alvarez, D. Franz, and G. Phillips for technical support.

## REFERENCES

1. **Almeida, R. P. P., and A. H. Purcell.** 2003. Biological traits of *Xylella fastidiosa* strains from grapes and almonds. Appl. Environ. Microbiol. **69:**7447–7452.
2. **Bhattacharyya, A., S. Stilwagen, G. Reznik, H. Feil, W. S. Feil, I. Anderson, A. Bernal, M. D'Souza, N. Ivanova, V. Kapatral, N. Larsen, T. Los, A. Lykidis, E. Selkov, Jr., T. L. Walunas, A. Purcell, R. A. Edwards, T. Hawkins, R. Haselkorn, R. Overbeek, N. C. Kyrpides, and P. F. Predki.** 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. Genome Res. **12:**1556–1563.
3. **Bichara, M., J. Wagner, and I. B. Lambert.** 2006. Mechanisms of tandem repeat instability in bacteria. Mutat. Res. **25:**144–163.
4. **Boles, B. R., M. Thoendel, and P. K. Singh.** 2004. Self-generated diversity produces "insurance effects" in biofilm communities. Proc. Natl. Acad. Sci. USA **101:**16630–16635.
5. **Chen, J., R. Groves, E. L. Civerolo, M. Viveros, M. Freeman, and Y. Zheng.** 2005. Two *Xylella fastidiosa* genotypes associated with almond leaf scorch disease on the same location in California. Phytopathology **95:**708–714.
6. **Coletta-Filho, H. D., M. A. Takita, A. A. de Souza, C. I. Aguilar-Vildoso, and M. A. Machado.** 2001. Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. Appl. Environ. Microbiol. **67:** 4091–4095.
7. **Collmer, A., and N. T. Keen.** 1986. The role of pectic enzymes in plant pathogenesis. Annu. Rev. Phytopathol. **24:**383–409.
8. **Davis, M. J., S. V. Thomson, and A. H. Purcell.** 1980. Etiological role of a xylem-limited bacterium causing Pierce's disease in almond leaf scorch. Phytopathology **70:**472–475.
9. **Davis, M. J., W. J. French, and N. W. Schaad.** 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. Curr. Microbiol. **6:**309–314.
10. **de Castro, L. A., T. Rodrigues Pedroso, S. S. Kuchiishi, M. Ramenzoni, J. D. Kich, A. Zaha, M. Henning Vainstein, and H. Bunselmeyer Ferreira.** 2006. Variable number of tandem amino acid repeats in adhesion-related CDS products in Mycoplasma hyopneumoniae strains. Vet. Microbiol. **116:**258–269.
11. **Fry, S. M., J. S. Huang, and R. D. Milholland.** 1994. Isolation and preliminary characterization of extracellular proteases produced by strains of *Xylella fastidiosa* from grapevines. Phytopathology **84:**357–363.
12. **Gartemann, K. H., O. Kirchner, J. Engemann, I. Gräfen, R. Eichenlaub, and A. Burger.** 2003. *Clavibacter michiganensis* subsp. *michiganensis*: first steps in the understanding of virulence of a gram-positive phytopathogenic bacterium. Bio/Technology **106:**179–191.
13. **Gravekamp, C., D. S. Horensky, J. L. Michel, and L. C. Madoff.** 1996. Variation in repeat number within the alpha C protein of group B streptococci alters antigenicity and protective epitopes. Infect. Immun. **64:**3576–3583.
14. **Hendson, M., A. H. Purcell, D. Chen, C. Smart, M. Guilhabert, and B. Kirkpatrick.** 2001. Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. Appl. Environ. Microbiol. **67:**895–903.
15. **Ho, Y. R., C. M. Li, H. P. Su, J. H. Wu, Y. C. Tseng, Y. J. Lin, and J. J. Wu.** 2007. Variation in the number of tandem repeats and profile of surface protein genes among invasive group B streptococci correlates with patient age. J. Clin. Microbiol. **45:**1634–1636.
16. **Jha, A. K., H. P. Bais, and J. M. Vivanco.** 2005. *Enterococcus faecalis* mammalian virulence-related factors exhibit potent pathogenicity in the *Arabidopsis thaliana* plant model. Infect. Immun. **73:**464–475.
17. **Jordan, P., L. A. Snyder, and N. J. Saunders.** 2003. Diversity in coding tandem repeats in related *Neisseria* spp. BMC Microbiol. **3:**23.
18. **Madoff, L. C., J. L. Michel, E. W. Gong, D. E. Kling, and D. L. Kasper.** 1996. Group B streptococci escape host immunity by deletion of tandem repeat elements of the alpha C protein. Proc. Natl. Acad. Sci. USA **93:**4131–4136.
19. **Maria Fedatto, L., M. E. Silva-Stenico, A. Etchegaray, F. T. Pacheco, J. L. Rodrigues, and S. M. Tsai.** 2006. Detection and characterization of protease secreted by the plant pathogen *Xylella fastidiosa*. Microbiol. Res. **161:**263–272.
20. **Mircetich, S. M., S. K. Lowe, W. J. Moller, and G. Nyland.** 1976. Etiology of almond leaf scorch disease and transmission of the causal agent. Phytopathology **66:**17–24.
21. **Nogaroto, V., S. Tagliavini, A. Gianotti, A. Mikawa, N. T. Barros, L. Puzer, A. K. Carmona, P. Costa, and F. Henrique-Silva.** 2006. Recombinant expression and characterization of a *Xylella fastidiosa* cysteine protease differentially expressed in a nonpathogenic strain. FEMS Microbiol. Lett. **261:** 187–193.
22. **Roper, M. C., L. C. Greve, J. G. Warren, J. M. Labavitch, and B. C. Kirkpatrick.** 2007. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Mol. Plant-Microbe Interact. **20:**411–419.
23. **Rozen, S., and H. J. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132:**365–386.
24. **Simpson, A. J., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, E. D. Bonaccorsi, S. Bordin, J. M. Bove, M. R. Briones, M. R. Bueno, A. A. Camargo, L. E. Camargo, D. M. Carraro, H. Carrer, N. B. Colauto, C. Colombo, F. F. Costa, M. C. Costa, C. M. Costa-Neto, L. L. Coutinho, M. Cristofani, E. Dias-Neto, C. Docena, H. El-Dorry, A. P. Facincani, A. J. Ferreira, V. C. Ferreira, J. A. Ferro, J. S. Fraga, S. C. Franca, M. C. Franco, M. Frohme, L. R. Furlan, M. Garnier, G. H. Goldman, M. H. Goldman, S. L.**

**Gomes, A. Gruber, P. L. Ho, J. D. Hoheisel, M. L. Junqueira, E. L. Kemper, J. P. Kitajima, J. E. Krieger, E. E. Kuramae, F. Laigret, M. R. Lambais, L. C. Leite, E. G. Lemos, M. V. Lemos, S. A. Lopes, C. R. Lopes, J. A. Machado, M. A. Machado, A. M. Madeira, H. M. Madeira, C. L. Marino, M. V. Marques, E. A. Martins, E. M. Martins, A. Y. Matsukuma, C. F. Menck, E. C. Miracca, C. Y. Miyaki, C. B. Monteriro-Vitorello, D. H. Moon, M. A. Nagai, A. L. Nascimento, L. E. Netto, A. Nhani, Jr., F. G. Nobrega, L. R. Nunes, M. A. Oliveira, M. C. de Oliveira, R. C. de Oliveira, D. A. Palmieri, A. Paris, B. R. Peixoto, G. A. Pereira, H. A. Pereira, Jr., J. B. Pesquero, R. B. Quaggio, P. G. Roberto, V. Rodrigues, A. J. D. Rosa, V. E. de Rosa, Jr., R. G. de Sa, R. V. Santelli, H. E. Sawasaki, A. C. da Silva, A. M. da Silva, F. R. da Silva, W. A. da Silva, Jr., J. F. da Silveira, M. L. Silvestri, W. J. Siqueira, A. A. de Souza, A. P. de Souza, M. F. Terenzi, D. Truffi, S. M. Tsai, M. H. Tsuhako, H. Vallada, M. A. Van Sluys, S. Verjovski-Almeida, A. L. Vettore, M. A. Zago, M. Zatz, J. Meidanis, and J. C. Setubal.** 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature **406:**151–157.

25. **Strand, M., T. A. Prolla, R. M. Liskay, and T. D. Petes.** 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature **365:**274–276.
26. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
27. **van Belkum, A., S. Scherer, L. van Alphen, and H. Verbrugh.** 1998. Short-sequence DNA repeats in prokaryotic genomes. Microbiol. Mol. Biol. Rev. **62:**275–293.
28. **van Belkum, A.** 1999. The role of short sequence repeats in epidemiologic typing. Curr. Opin. Microbiol. **2:**306–311.
29. **Van Sluys, M. A., M. C. de Oliveira, C. B. Monteiro-Vitorello, C. Y. Miyaki, L. R. Furlan, L. E. Camargo, A. C. da Silva, D. H. Moon, M. A. Takita, E. G. Lemos, M. A. Machado, M. I. Ferro, F. R. da Silva, M. H. Goldman, G. H. Goldman, M. V. Lemos, H. El-Dorry, S. M. Tsai, H. Carrer, D. M. Carraro, R. C. de Oliveiram, L. R. Nunes, W. J. Siqueira, L. L. Coutinho, E. T. Kimura, E. S. Ferro, R. Harakava, E. E. Kuramae, C. L. Marino, E. Giglioti, I. L. Abreu, L. M. Alves, A. M. do Amaral, G. S. Baia, S. R. Blanco, M. S. Brito, F. S. Cannavan, A. V. Celestino, A. F. da Cunha, R. C. Fenille, J. A. Ferro, E. F. Formighieri, L. T. Kishi, S. G. Leoni, A. R. Oliveira, V. E. Rosa, Jr., F. T. Sassaki, J. A. Sena, A. A. de Souza, D. Truffi, F. Tsukumo, G. M. Yanai, L. G. Zaros, E. L. Civerolo, A. J. Simpson, N. F. Almeida, Jr., J. C. Setubal, and J. P. Kitajima.** 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J. Bacteriol. **185:**1018–1026.
30. **Verstrepen, K. J., A. Jansen, F. Lewitter, and G. R. Fink.** 2005. Intragenic tandem repeats generate functional variability. Nat. Genet. **37:**986–990.
31. **Wells, J. M., B. C. Raju, H. Y. Hung, W. G. Weisburg, L. Mandelco-Paul, and B. J. Brenner.** 1987. *Xylella fastidiosa* new-genus new-species gram-negative xylem-limited fastidious plant bacteria related to *Xanthamonas* spp. Int. J. Syst. Bacteriol. **37:**136–143.
32. **Wulff, N. A., H. Carrer, and S. F. Pascholati.** 2006. Expression and purification of cellulase Xf818 from *Xylella fastidiosa* in *Escherichia coli*. Curr. Microbiol. **53:**198–203.