Development of Tools to Detect and Identify Strains Belonging to the Pseudomonas syringae Species Complex Responsible for Vein Clearing of Zucchini

Caroline Lacault, Martial Briand, Marie-Agnès Jacques, Armelle Darrasse

PhytoFrontiers™
An Open Access Journal from The American Phytopathological Society

Research

Development of Tools to Detect and Identify Strains Belonging to the *Pseudomonas syringae* Species Complex Responsible for Vein Clearing of Zucchini

Caroline Lacault | Martial Briand | Marie-Agnès Jacques | Armelle Darrasse†

Univ Angers, Institut Agro, INRAE, IRHS, SFR QUASAV, Angers, F-49000, France

Accepted for publication 6 May 2024.

Abstract

Vein clearing of zucchini (VCZ) is a disease caused by seedborne bacteria that affects young plants of *Cucurbita pepo* subsp. *pepo*. VCZ agents are distributed into four phylogenetic clusters within clades 2a and 2ba of phylogroup 2 of the *Pseudomonas syringae* species complex. All these strains are pathogenic to squash, but only certain strains can also attack melon and cucumber. Strains belonging to clades 2b and 2d are sometimes isolated from zucchini seeds but have not been associated with VCZ epidemics. Identification tools for VCZ agents are required to improve disease control. Primers were designed to implement a seven-gene multilocus sequence analysis (MLSA) scheme for a collection of 60 strains isolated from zucchini seeds. The MLSA showed a clear predominance of strains of cluster 2ba-A and the presence of VCZ strains in a fifth cluster (2ba-C). PCR tests were designed to characterize strains in the VCZ clusters, and a multiplex qPCR test was proposed to distinguish strains with a cucurbit host range extended to melon and cucumber, harboring *hopZ5* and *sylC*, from other strains harboring *avrRpt2* and *sylC*. Additional qPCR tests were also designed to gain insights into clade-2b and -2d strains that can be isolated from cucurbits. These tools evaluated in silico with the NCBI database and experimentally with a collection of 112 strains detected all target strains, except for the test dedicated to clade-2b strains, and excluded 96.7 to 100% of nontarget strains. These tools are intended to serve phylogenetic studies, epidemiological monitoring, and seed testing.

Keywords: *Cucurbita pepo*, detection test, diagnosis, MLSA, qPCR, seed, seedborne pathogen, TaqMan

† Corresponding author: A. Darrasse; armelle.darrasse@inrae.fr

e-Xtra: Supplementary material is available online.

The author(s) declare no conflict of interest.

Vein clearing (VCZ) is a bacterial disease that affects young plants of zucchini (*Cucurbita pepo* subsp. *pepo*), leading to plantlet deformations, water-soaked and necrotic spots on cotyledons and leaves, and stunting (Lacault et al. 2020). A transient symptom of vein clearing may be observed on the first leaves, but no symptoms have been

reported so far on adult field-grown plants (Lacault et al. 2020; Manceau et al. 2011). VCZ is caused by seed-transmitted bacterial pathogens, as a clear link has been established between sowing infected seed lots and disease outbreaks in nurseries, and seed infection negatively impacts germination (Manceau et al. 2011).

The causal agents of VCZ belong to the *Pseudomonas syringae* species complex (Lacault et al. 2020) and are closely related to strains that cause symptoms on adult cucurbit plants, such as systemic infection of zucchini (Djitro et al. 2022b) and bacterial leaf spots (Newberry et al. 2019; Qiao et al. 2023). The *P. syringae* species complex (Gomila et al. 2017) includes bacteria that can cause severe epidemics on woody plants and annual crops (for reviews, see Lamichhane et al. 2014, 2015) but also many strains that are present in the environment, some of which are phytopathogenic (Morris et al. 2013). The taxonomy of the *P. syringae* species complex has been revised with the description of 13 phylogroups (PGs) subdivided into clades (Berge et al. 2014), and 19 phylogenomic species proposed (Gomila et al. 2017). Comparative genomics of strain collections has improved the knowledge of the structure and evolution of this species complex, highlighting impacts of recombination and horizontal gene transfer events. Homologous recombination blurs boundaries of some species (Dillon et al. 2019b). Based on former studies that identified and characterized different functions involved in pathogenicity (i.e., toxins, type three secretion systems [T3SSs], and repertoires of type three effectors [T3Es]), comparative genomics allows discrimination of bacteria with different behaviors. For example, the presence in a genome sequence of a canonical T3SS, numerous T3Es, and few toxins is associated with a pathogenic bacterium well adapted to a restricted number of host plants. In contrast, few T3Es and more toxins correspond to a more generalist pathogen, and when a genome sequence also lacks the canonical T3SS, this bacterial strain is mostly nonpathogenic (Baltrus et al. 2017; Dillon et al. 2019a; Xin et al. 2018).

VCZ strains are diverse genetically as the genome sequences of these strains are distributed into four clusters of the *P. syringae* species complex (Lacault et al. 2020). Clusters 2ba-A and 2ba-B belong to clade 2ba (Lacault et al. 2020), which resulted from homologous recombination between strains in clades 2a and 2b (Newberry et al. 2019). Strains in clusters 2a-D and 2a-E belong to clade 2a (Lacault et al. 2020), corresponding with strains of *P. cerasi* (Gomila et al. 2017). These clusters also encompass strains that are pathogenic on other cucurbits and share average nucleotide identities based on BLAST (ANIb) values with VCZ strains as high as those observed between two VCZ strains (Lacault et al. 2020; Newberry et al. 2019). Two closely related clusters, clusters 2ba-C and 2a-F, group strains that have been sampled from watermelon (*Citrullus lanatus*), cantaloupe (*Cucumis melo*), and squash (*Cucurbita* spp.) in the United States (Newberry et al. 2019). Recently, clusters 2ba-A and 2ba-C strains causing unusual symptoms on zucchini, such as twisted petioles, necrotic leaves, crown rots, and internal fruit rots have been reported in Australia (Djitro et al. 2022a, b). An extended multilocus sequence analysis (MLSA) scheme has been proposed to allocate strains in these clusters and facilitate the analysis of large strain collections (Lacault et al. 2020). It includes the previously available *P. syringae* scheme based on four housekeeping genes (*gapA*, *gltA*, *gyrB*, and *rpoB*; Hwang et al. 2005), and three widely distributed genes (*Psyr3420*, *Psyr4880*, and *Psyr3208*).

The various VCZ clusters of strains have different host ranges. VCZ strains from cluster 2ba-A are characterized by a narrow host range of cucurbits that is limited to the genus *Cucurbita*. All other VCZ strains have a wider cucurbit host range, affecting squash, melon, cucumber (*Cucumis sativus*), and sometimes watermelon (Lacault et al. 2020). The extent of these host ranges is consistent with their repertoire of T3Es. In particular, cluster 2ba-A strains, with the narrowest host range, are characterized by the presence of *avrRpt2*, while the clusters with a wider host range harbor *hopZ5*. Similarly, Djitro et al. (2022a) reported that Australian strains harboring *avrRpt2* (from cluster 2ba-A and part of cluster 2ba-C, proposed to be named 2ba-C2) have a more limited host range on cucurbits than other 2ba-C strains harboring *hopZ5* (proposed to be named 2ba-C1).

Strains pathogenic on cucurbits are also identified in clades 2b and 2d, corresponding to the *P. syringae* species *sensu stricto* and to the new genomic species A (Gomila et al. 2017), respectively. One strain, isolated from a zucchini seed lot and causing symptoms on zucchini, is close genetically to clade-2b strains belonging to the pathovars *aptata* and *syringae* (Lacault et al. 2020), which are reported to be pathogenic on beet (*Beta vulgaris*) and lilac (*Syringa vulgaris*), respectively (Lamichhane et al. 2015), and can cause diseases on cucurbits such as cantaloupe and squash (Langston et al. 2003; Morris et al. 2000; Sedighian et al. 2014). Clade-2d strains isolated from watermelon and cantaloupe plants in the United States and France were rather aggressive on their isolation hosts (Berge et al. 2014; Newberry et al. 2016) but weakly aggressive on squash (Newberry et al. 2016).

Seed testing is a first step to control seedborne pathogenic bacteria, and it requires specific and sensitive detection tools to detect the pathogenic bacteria among a variety of microorganisms, including look-alikes or closely related strains that are not pathogenic, but which also can be seedborne (Jacques et al. 2012). Currently, seed-associated plant pathogenic bacteria are mainly detected using PCR-based tests, due to the ease of use and typical specificity of these assays. A large range of PCR-based detection and identification tools targeting different phylogenetic groups of the *P. syringae* species complex have been developed (Borschinger et al. 2016; Guilbaud et al. 2016), such as PCR tests targeting genes encoding toxins and other pathogenicity factors (Bereswill et al. 1994; Bultreys and Gheysen 1999; Chen et al. 2020; Gallelli et al. 2014; Meng et al. 2017; Prosen et al. 1993). A qPCR test targeting *sylC* (4Ba marker) was designed to detect VCZ strains (Manceau et al. 2011). Indeed, *sylC* is one of the genes encoding the nonribosomal peptide synthetase responsible for syringolin synthesis (Amrein et al. 2004). The syringolin biosynthesis genes are widely distributed in phylogroup 2 (PG2) strains (Baltrus et al. 2011; Dillon et al. 2019b; Dudnik and Dudler 2014; Hulin et al. 2018). This marker may, hence, not be specific enough for VCZ strain detection.

The aim of the study was to design and evaluate new identification and detection tools for VCZ strains. First, we developed PCR tests to amplify three genes to be included in the seven-gene MLSA scheme and characterized our collection of 53 VCZ strains with this scheme. Second, we gained insight into clade-2d strains isolated from zucchini seeds, for which very little data were available, by sequencing two genomes. Third, faced with the genetic diversity of the strains responsible for VCZ, we used two approaches to design identification and detection tests. The first approach was based on the search for phylogenetic markers for the clusters and phylogenetic groups to which the VCZ strains belonged. The second was to design a multiplex TaqMan qPCR assay associating *sylC* and other pathogenicity markers such as *avrRpt2* and *hopZ5*, whose presence characterized VCZ strains according to their host range on cucurbits. Finally, the multiplex qPCR test was evaluated for detection of VCZ strains in cucurbit seed samples.

Materials and Methods

Bacterial strains and growth

This study used a bacterial collection of 112 strains, including 54 strains responsible for VCZ (Lacault et al. 2020), 14 strains isolated from zucchini seeds or seed-production fields and belonging to clade 2b (six strains) and clade 2d (eight strains), 35 strains from diverse *Pseudomonas* species and pathovars, 7 bacterial strains pathogenic on cucurbits belonging to divergent genera (*Pectobacterium*, *Erwinia*, *Ralstonia*, *Xanthomonas*, and *Acidovorax*), and 2 strains associated with cucurbit seeds belonging to *Bacillus* and *Paenibacillus* (Khalaf and Raizada 2016) (Table 1). Representatives of other genera, species, and pathovars were provided by the French Collection of Plant-Associated Bacteria (https://cirm-cfbp.fr/).

Bacterial strains were stored in a −80°C freezer in 40% glycerol solution and were cultivated routinely at 28°C on 10% TSA medium (3 g liter^{-1} of tryptone soya broth and 15 g liter^{-1} of agar) with 50 mg liter^{-1} of cycloheximide. Bacterial suspensions were prepared from fresh cultures suspended in sterile distilled water, calibrated at $OD_{650\,nm} = 0.1$ (around 1×10^8 CFU ml^{-1}), and adjusted to the final desired concentration with sterile distilled water. Dilution plating and colony counts after 48-h growth on 10% TSA were used to quantify bacterial population sizes precisely. For PCR and qPCR assays, bacterial suspensions were boiled for 10 min before use.

Amplification and sequencing of the three genes to be included in the MLSA scheme

Three genes have been described for typing VCZ strains (Lacault et al. 2020). In the genome sequence of strain B728a, these genes correspond to (i) *Psyr3208*, encoding the NADH dehydrogenase subunit M; (ii) *Psyr3420*, encoding a protein belonging to the FixH superfamily; and (iii) *Psyr4880*, encoding a conserved hypothetical protein with unknown function. The sequence of each gene was retrieved from 118 genome sequences, including those of 23 strains isolated from zucchini (Lacault et al. 2020), 14 strains isolated from various cucurbits (Newberry et al. 2019), the 2 clade-2d strains isolated from zucchini seeds (this work), and 79 other PG2 strains (Supplementary Table S1), and aligned. Primers were designed based on the conserved regions in order to amplify and sequence polymorphic fragments. The primer pairs Psyr3208-F 5′-ATGATGCTGGTGCCGATGTA-3′ and Psyr3208-R 5′-ATCAGCGAGTAGACAGAACC-3′, Psyr3420-F 5′-ATGGTGTTCATCGCCGTGACC-3′ and Psyr3420-R 5′-TCGCCTTCGACACCCAGCAG-3′, and Psyr4880-F 5′-ATGCACGGCTCCGAACGAAA-3′ and Psyr4880-R 5′-AGTCTCCAGCGCCCGTCCAC-3′ resulted in the amplification of 883-, 343-, and 416-bp fragments, respectively. To increase the length of double-stranded sequences, supplementary primers Psyr3208-MF 5′-TCGGCCGAGTTCGCACCGAT-3′ and Psyr3208-MR 5′-CTGCCCGAGTAGATACCGAT-3′ were designed in the middle of the *Psyr3208* fragment to be used with Psyr3208-R and Psyr3208-F, respectively. PCR assays were performed in a 50-µl volume containing 10 µl of buffer, 0.4 U µl^{-1} of Taq polymerase (GoTaq, Promega, Charbonnières-les-Bains, France), 200 µM of dNTPs, 0.5 µM of each primer, and 5 µl of boiled bacterial suspension. For all *Psyr3208* primers, amplification conditions were 5 min at 94°C followed by 20 cycles of 30 s at 94°C, 30 s at 60°C (with a 0.5°C decrease at each cycle), and 1 min at 72°C, then 15 cycles of 30 s at 94°C, 30 s at 50°C, and 1 min at 72°C. For *Psyr3420* or *Psyr4880* primers, PCR conditions were 5 min at 94°C followed by 35 cycles of 30 s at 94°C, 30 s at 62°C (for *Psyr3420*) or 66°C (for *Psyr4880*), and 1 min at 72°C. A final step of 10 min at 72°C was added to achieve elongation for all PCR tests. Amplified fragments were Sanger sequenced with the forward and reverse primers, and with MF- and MR-primers for *Psyr3208* (Genoscreen, Lille, France). Geneious version 9.1.7 software was used to assemble, align, orientate, and trim sequences according to the reading frame. Sequences of partial coding DNA sequences (CDSs) were deposited in GenBank under the accession numbers MW892843 to MW892968, and were 825- (*Psyr3208*), 318- (*Psyr3420*), and 309-bp (*Psyr4880*) long. Genes were amplified from strains for which the genome sequences were not available. Homologous sequences retrieved from genome sequences (Table 2) were added to the data set.

MLSA of the strain collection

The sequences of *gapA*, *gltA*, *gyrB*, and *rpoD* (Lacault et al. 2020), encoding glyceraldehyde-3-phosphate dehydrogenase A, citrate synthase A, B subunit of DNA gyrase, and RNA polymerase sigma70 factor RpoD, respectively, were concatenated to the sequences of *Psyr3208*, *Psyr3420*, and *Psyr4880* to obtain a 3,459-bp fragment for each strain. A maximum likelihood tree was generated with MEGA 7 (Kumar et al. 2016). The best fit model was the Tamura-Nei model using the discrete Gamma distribution with invariant sites (G + I), and 1,000 bootstrap replicates were used.

Genome sequencing and analyses

To design identification tools based on genome sequences of clade-2d strains isolated from cucurbits, the genomes of strains $P129_{2d}$ and $P135_{2d}$ were sequenced. DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega) and was paired-end sequenced (2 × 150 bp) using the Illumina HiSeq X-Ten at Beijing Genomics Institute (Shenzhen, China). Genome assembly was performed with SOAPdenovo version 2.04, SOAPGapCloser version 1.12 (Luo et al. 2012), and Velvet version 1.2.02 (Zerbino and Birney 2008). Fasta-formatted assemblies were deposited in NCBI with accession numbers JAGPYS000000000 and JAGPYR000000000. The percentage of shared k-mers and distance matrix were calculated based on the dataset of 118 genome sequences (Supplementary Table S1) with Ki-S, as were the dendrograms (https://iris.angers.inra.fr/galaxypub-cfbp) (Briand et al. 2021) to determine the phylogenetic position of clade-2d strains.

The T3E repertoire was predicted in each genomic sequence through a tBLASTN search (identity ≥ 80% for at least 80% of the CDS length) for more than 700 amino acid sequences of T3Es, according to the Hopsdatabase (https://web.archive.org/web/20240216130544/www.pseudomonas-syringae.org/).

Spiking of seed samples, sampling seed lots, and DNA extraction for qPCR assays

Healthy seed samples (negative result for VCZ strains of bacteria based on tests of 2,000 seeds) of *C. pepo* subsp. *pepo* cultivar Lorea (500 seeds) were soaked in 3 ml g^{-1} of seeds of phosphate-buffered saline (PBS, Sigma-Aldrich, Schnelldorf, Germany) plus Tween 20 (0.02% vol/vol) and shaken (105 rpm) for 2 h at room temperature. A fresh bacterial suspension calibrated at 1×10^8 CFU ml^{-1} was diluted with each seed extract to produce spiked suspensions at concentrations from 1 to 1×10^7 CFU ml^{-1}, as confirmed by dilution plating. One ml per dilution was centrifuged for 5 min at 13,000 × *g*. The supernatant was discarded and DNA extracted using the QuicPick SML Plant DNA purification kit according to the supplier's instructions (BioNobile, Turku, Finland) in an automated system (Caliper Zephyr, PerkinElmer, Villebon-sur-Yvette, France). The DNA then was eluted from the magnetic beads using 30 µl of elution buffer. One

TABLE 1

Results of multilocus sequence analysis (MLSA), PCR, and qPCR assays on a collection of *Pseudomonas syringae* strains responsible for vein clearing of zucchini (VCZ) and other bacterial strains

			Identification by PCR assay[d]											
			2ba				2a			Identification by qPCR assay[d,e]				
Strain code[a]	Strain information[b]	MLSA group cluster[c]	A	B	C	ABC	D	E	DE	4Ba	Rpt2	Z5	2d	2b*
Strains from main VCZ clusters isolated from zucchini seeds (Lacault et al. 2020)														
P99	France, 2005	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12832	France, 2006	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12854	France, 2002	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12855	Chile, 2007	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P113	Thailand, 2011	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P118	France, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P121	Chile, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P123	China, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P127	India, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P139	China, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P22	U.S.A., 2009	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P5	France, 2008	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P79	France, 2010	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P84	China, 2010	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12831	France, 2005	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P112	Chile, 2011	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P115[f]	China, 2011	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P119	France, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12	France, 2006	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P120	France, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P125	China, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12833	France, 2005	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P12856	France, 2005	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P130	Thailand, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P131	China, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P133	Thailand, 2013	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P137	France, 2014	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P24	France, 2008	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P4	France, 2005	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P68	U.S.A., 2010	2ba-A	+	−	−	+	−	−	−	+	+	−	−	−
P55	Chile, 2009	2ba-B	+	+	−	+	−	−	−	+	+	−	−	−
P56	na, 2009	2ba-B	+	+	−	+	−	−	−	+	+	−	−	−
P12857	France, 2007	2ba-B	−	+	−	+	−	−	−	+	−	+	−	−
P1	France, 2005	2ba-B	−	+	−	+	−	−	−	+	−	+	−	−
P12836	France, 2005	2ba-B	−	+	−	+	−	−	−	+	−	+	−	−
P14	na, 2008	2ba-B	−	+	−	+	−	−	−	+	−	+	−	−
P69	U.S.A., 2010	2ba-B	−	+	−	+	−	−	−	+	−	+	−	−
P88	China, 2011	2ba-C	−	−	+	+	−	−	−	+	−	+	−	−
P100	na, 2011	2ba-C	−	−	+	+	−	−	−	+	−	+	−	−
P77	France, 2010	2a-D	−	−	−	−	+	−	+	+	−	+	−	−
P62	France, 2010	2a-D	−	−	−	−	+	−	+	+	−	+	−	−
P70	Thailand, 2010	2a-D	−	−	−	−	+	−	+	+	−	+	−	−
P71	na, 2010	2a-D	−	−	−	−	+	−	+	+	−	+	−	−
P72	na, 2010	2a-D	−	−	−	−	+	−	+	+	−	+	−	−
P66	Thailand, 2010	2a-E	−	−	−	−	−	+	+	+	−	+	−	−
P65	na, 2010	2a -E	−	−	−	−	−	+	+	+	−	+	−	−
P89	France, 2011	2a-E	−	−	−	−	−	+	+	+	−	+	−	−
P90	U.S.A., 2011	2a-E	−	−	−	−	−	+	+	+	−	+	−	−
P94	na, 2011	2a-E	−	−	−	−	−	+	+	+	−	+	−	−
P104	na, 2011	2a-E	−	−	−	−	−	+	+	+	−	+	−	−
Other strains isolated from zucchini seed lots or seed production fields														
P108	Seed, Chile, 2011	2b	−	−	−	−	−	−	−	+	−	−	−	+
P101	Seed, 2011	2b	−	−	−	−	−	−	−	+	−	−	−	+
P136[f]	Seed, India, 2014	2b	−	−	−	−	−	−	−	+	−	−	−	+
P138	Seed, France, 2015	2b	−	−	−	−	−	−	−	+	−	−	−	+

(*Continued on next page*)

[a] T = type strain and PT = pathotype strain.

[b] na = not available; within the *P. syringae* complex, the species name is based on Gomila et al. (2017).

[c] Phylogroups and clades according to Berge et al. (2014) and Newberry et al. (2019); clusters according to Lacault et al. (2020) and to the new MLSA scheme developed in this study that includes seven genes.

[d] + = positive result with PCR test; − = negative result with PCR test; w = presence of a band of the wrong size.

[e] Identification by the multiplex and simplex qPCR assays using the pathogenicity makers 4Ba (*sylC*), Rpt2 (*avrRpt2*), and Z5 (*hopZ5*), and phylogenetic marker 2d for clade-2d strains, and in simplex qPCR assay for phylogenetic marker 2b* for strains belonging to a subgroup of clade 2b that did not encompass strains of clade 2ba or the pathovars *pisi*, *coryli*, and *avellanae*.

[f] These strains were not included in the compete MLSA scheme.

[g] Strain CFBP 1750 is synonymous with strain NCPPB 4996. Based on its name, this strain should belong to phylogroup 4; however, it is classified in clade 2b according to the closest genome sequences identified in the NCBI database using KISonWGS and Dillon et al. (2019b).

to five µl of DNA were used per qPCR sample. Boiled bacterial suspensions in sterile distilled water, ranging from 1 to 1 × 10^7 CFU ml^{-1}, were also used as standards for quantification. Each sample was tested in triplicate. Four commercial zucchini seed lots (I to IV) were each tested using three subsamples of 100 seeds. Depending on the number of positive subsamples, additional subsamples of 500, 100, or 10 seeds were tested to obtain subsamples contaminated and uncontaminated with VCZ strains of bacteria. The number of contaminated or uncontaminated subsamples enabled determination of the contamination rate of the seed lot as described by Maury et al. (1986) for subsamples of the same size, or according to Swaroop (1951) for subsamples of different sizes. One µl of DNA was used per qPCR sample.

Identification of target sequences to be used for the design of specific PCR assays

Specific sequences of bacterial strains in clades and clusters were searched in the genomic sequences using the SkIf tool (Briand et al. 2016; https://iris.angers.inra.fr/galaxypub-cfbp SkIf_with_DSK). SkIf_with_DSK identifies specific k-mers within a group of genome sequences (in-group) that are absent in other genome sequences (out-group). It also provides their

TABLE 1 (*Continued from previous page*)

			Identification by PCR assay[d]							Identification by qPCR assay[d,e]				
			2ba				2a							
Strain code[a]	Strain information[b]	MLSA group cluster[c]	A	B	C	ABC	D	E	DE	4Ba	Rpt2	Z5	2d	2b*
107[f]	Leaf, France, 2018	2b	–	–	w	–	–	–	–	+	–	–	–	+
598p[f]	Leaf, France, 2018	2b	–	–	–	–	–	–	–	+	–	–	–	+
P129	Seed, China, 2013	2d	–	–	–	–	–	–	–	+	–	–	+	–
P135	Seed, India, 2014	2d	–	–	–	–	–	–	–	+	–	–	+	–
P103	Seed, 2011	2d	–	–	–	–	–	–	–	+	–	–	+	–
P134	Seed, 2014	2d	–	–	–	–	–	–	–	+	–	–	+	–
84[f]	*Beta vulgaris*, France, 2018	2d	–	–	–	–	–	–	–	+	–	–	+	–
142[f]	Seed, France, 2018	2d	–	–	–	–	–	–	–	+	–	–	+	–
475[f]	Seed, France, 2018	2d	–	–	–	–	–	–	–	+	–	–	+	–
485[f]	*Solanum nigrum*, France, 2018	2d	–	–	–	–	–	–	–	+	–	–	+	–
Other strains of the *Pseudomonas* species complex														
CFBP 2215	*P. syringae* pv. *delphinii*	1a	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 1727	*P. syringae* pv. *berberidis*	1a	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 1657[PT]	*P. syringae* pv. *maculicola*	1a	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 5524[PT]	*P. syringae* pv. *spinaceae*	1a	–	–	–	–	–	–	–	–	+	–	–	–
CFBP 2212[PT]	*P. tomato*	1a	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 1698	*P. tomato*	1a	–	–	–	–	–	–	–	–	+	–	–	–
CFBP 4060[T]	*P. avellanae*	1b	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 7286	*P. avellanae* pv. *actinidiae*	1b	–	–	–	–	–	–	–	–	–	+	–	–
CFBP 8305[T]	*P. cerasi*	2a	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 2356[PT]	*P. cerasi* pv. *dysoxyli*	2a	–	–	–	–	–	–	–	+	–	–	–	–
CFBP 1392[T]	*P. syringae* pv. *syringae*	2b	–	–	–	–	–	–	–	+	–	–	–	+
CFBP 1617[PT]	*P. syringae* pv. *aptata*	2b	–	–	–	–	–	–	–	+	–	–	–	+
CFBP 7019[T]	*P. congelans*	2c	–	–	–	–	–	–	–	–	–	–	–	–
CC94	*P. syringae* pv. *aptata*	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 2339[PT]	Species A (pv. *aceris*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
B728A	Species A (pv. *syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 8529	Species A (pv. *syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 8552	Species A (*P. syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 8583	Species A (*P. syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 7148	Species A (pv. *solidagae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 8553	Species A (pv. *syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 8501	Species A (pv. *syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 8538	Species A (*P. syringae*)	2d	–	–	–	–	–	–	–	+	–	–	+	–
CFBP 2894	*P. amygdali* pv. *aesculi*	3	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 6463[PT]	*P. amygdali* pv. *lachrymans*	3	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 3651	*P. amygdali* pv. *phaseolicola*	3	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 1390[PT]	*P. amygdali* pv. *phaseolicola*	3	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 3205[T]	*P. amygdali*	3	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 2216[T]	*P. coronafaciens*	4	–	+	–	–	–	–	–	–	–	–	–	–
CFBP 1750	*coronafaciens* pv. *striafaciens*	2b[g]	–	–	–	–	–	–	–	–	–	–	–	+
CFBP 2341[T]	*P. cannabina*	5	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 3204[T]	*P. caripapayae*	6	–	–	–	–	–	–	–	–	–	–	–	–
Other cucurbit pathogenic strains														
CFBP 4459[T]	*Acidovorax citrulli*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 4375	*A. citrulli*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 2355[T]	*Erwinia tracheiphila*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 1350	*Pectobacterium carotovorum*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 6776	*Ralstonia solanacearum*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 2542	*Xanthomonas cucurbitae*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 4644	*X. melonis*	–	–	–	–	–	–	–	–	–	–	–	–	–
Strains representative of species found in zucchini seed microbiota														
CFBP 4246[T]	*Bacillus amyloliquefaciens*	–	–	–	–	–	–	–	–	–	–	–	–	–
CFBP 4258[T]	*Paenibacillus polymyxa*	–	–	–	–	–	–	–	–	–	–	–	–	–

TABLE 2

Genome sequences of strains belonging to the species complex *Pseudomonas syringae* used for multilocus sequence analysis (MLSA)

Strain code[a]	Country of isolation[b]	Host of isolation	Year of isolation[b]	Phylogroup, clade, and cluster[c]	Accession number[d]
ICMP 2844^{PT} (pv. *tomato*)	U.K.	*Lycopersicon esculentum*	1960	1a	LJRN00000000
Pmp 19322^{PT} (pv. *morsprunorum*)	U.S.A.	*Prunus domestica*	na	1b	RBNS00000000
ICMP 3923^{PT} (pv. *theae*)	Japan	*Thea sinensis*	1970	1b	LJRU00000000
Cit7 (*P. syringae*)	na	*Citrus sinensis*	na	2a	CP073636
58T^{T} (*P. cerasi*)	Poland	*Prunus cerasus*	2007	2a	GCA_900074915.1
ICMP 4048^{PT} (pv. *papulans*)	Canada	*Malus pumila*	1973	2a	LJRB00000000
03-19A	Florida, U.S.A.	*Cucumis melo*	2003	2a-D	MUHN00000000
200-1	Georgia, U.S.A.	*Cucurbita* sp.	2000	2a-D	MVAZ00000000
P77	France	*Cucurbita pepo*	2010	2a-D	WJPT00000000
P73	U.S.A.	*C. pepo*	2010	2a-D	WJPS00000000
P78	France	*C. pepo*	2010	2a-D	WJPU00000000
P87	China	*C. pepo*	2011	2a-D	WJPX00000000
BS2121	California, U.S.A.	*Cucurbita* sp.	2006	2a-E	MVAV00000000
ZUM3984	China	*Cucurbita* sp.	2008	2a-E	MVAX00000000
CFBP 8692 (P66)	Thailand	*C. pepo*	2010	2a-E	WJSF00000000
P89	France	*C. pepo*	2011	2a-E	WJPY00000000
13-139B	Florida, U.S.A.	*Citrullus lanatus*	2013	2a-F	MVAT00000000
13-429	Florida, U.S.A.	*C. lanatus*	2013	2a-F	MVAY00000000
ICMP 459^{PT} (pv. *aptata*)	U.S.A.	*Beta vulgaris*	1959	2b	LJRP00000000
ICMP 3023^{T} (pv. *syringae*)	U.K.	*Syringa vulgaris*	1951	2b	LJRK00000000
NCPPB 4273^{PT} (pv. *coryli*)	Italy	*Corylus avellana*	1995	2b	AWQP00000000
P108	Chile	*C. pepo*	2011	2b	VLIA00000000
1188_1	California, U.S.A.	*C. pepo*	na	2b	RBPG00000000
HS191 (pv. *syringae*)	Australia	*Panicum miliaceum*	1969	2ba	GCA_000988395.1
CFBP 8693 (P99)	France	*C. pepo*	2005	2ba-A	WJSE00000000
P12832	France	*C. pepo*	2006	2ba-A	VLHX00000000
P12854	France	*C. pepo*	2002	2ba-A	VLHY00000000
P12855	Chile	*C. pepo*	2007	2ba-A	VLHZ00000000
P113	Thailand	*C. pepo*	2011	2ba-A	VLIC00000000
P118	France	*C. pepo*	2013	2ba-A	VLID00000000
P121	Chile	*C. pepo*	2013	2ba-A	VLIE00000000
P123	China	*C. pepo*	2013	2ba-A	VLIF00000000
P127	India	*C. pepo*	2013	2ba-A	VLIG00000000
P139	China	*C. pepo*	2015	2ba-A	VLIH00000000
P22	U.S.A.	*C. pepo*	2009	2ba-A	WJPQ00000000
P5	France	*C. pepo*	2008	2ba-A	WJPR00000000
P79	France	*C. pepo*	2010	2ba-A	WJPV00000000
P84	China	*C. pepo*	2010	2ba-A	WJPW00000000
P12831 (co7A)	France	*C. pepo*	2005	2ba-A	WJPZ00000000
PS711	Serbia	*Cucurbita* sp.	2013	2ba-A	RQXZ00000000
13-509A	Florida, U.S.A.	*Cucurbita* sp.	2013	2ba-A	MUHP00000000
KL004-k1	Australia	*C. pepo*	2016	2ba-A	JAJITA00000000
13-C2	Florida, U.S.A.	*C. lanatus*	2013	2ba-B	MUHO00000000
13-140A	Florida, U.S.A.	*C. lanatus*	2013	2ba-B	MUHL00000000
14-410	Florida, U.S.A.	*C. lanatus*	2014	2ba-B	MUHQ00000000
14-32	Florida, U.S.A.	*C. lanatus*	2014	2ba-B	MUHM00000000
14-Gil	Florida, U.S.A.	*C. lanatus*	2014	2ba-B	MVAU00000000
CC457 (*P. syringae*)	France	*C. melo*	na	2ba-B	GCA_000452585.3
P12857	France	*C. pepo*	2007	2ba-B	VLIB00000000
ZUM3584	Italy	*Cucurbita* sp.	2005	2ba-C	MVBA00000000
KFR003-1	Australia	*C. pepo*	2016	2ba-C	JAJITB00000000
77-4C	Australia	*C. pepo*	2016	2ba-C	SRR21657462
ICMP 19117^{PT} (*P. congelans*)	Germany	Poaceae	1994	2c	LJQB00000000
B728A (pv. *syringae*)	U.S.A.	*Phaseolus vulgaris*	na	2d	QJTV00000000
ICMP 16925^{PT} (pv. *solidagae*)	Japan	*Solidago altissima*	na	2d	LJRH00000000
ICMP 2802^{PT} (pv. *aceris*)	U.S.A.	*Acer* sp.	na	2d	LJPM00000000
1212 (pv. *syringae*)	U.K.	*Pisum sativum*	na	2d	AVCR00000000
HRI-W 7872 (pv. *syringae*)	U.K.	*Prunus domestica*	2000	2d	LIHS00000000
P129	China	*C. pepo*	2013	2d	JAGPYS00000000
P135	India	*C. pepo*	2014	2d	JAGPYR00000000
ICMP 2740^{PT} (pv. *phaseolicola*)	Canada	*P. vulgaris*	1949	3	LJQZ00000000
NM002 (pv. *lachrymans*)	China	*Cucumis sativus*	1984	3	GCA_002068135.1
LMG 5060^{PT} (*P. coronafaciens*)	U.K.	*Avena sativa*	1958	4	JSED00000000
ICMP 2823^{PT} (pv. *cannabina*)	Hungary	*Cannabis sativus*	1957	5	FNKU00000000
ICMP 4091^{PT} (pv. *tagetis*)	Zimbabwe	*Tagetes erecta*	1972	6	LJRM00000000
ICMP 3272 (*P. viridiflava*)	New Zealand	*Actinidia deliciosa*	1971	8	RBQZ00000000

[a] ^{T} = type strain; ^{PT} = pathotype strain.

[b] na = not available.

[c] Phylogroup, clade, and cluster according to Berge et al. (2014), Newberry et al. (2019), and Lacault et al. (2020).

[d] When accession number was not available, assembly number was provided. When assembly was not available, SRA number was used. In the latter case, reads were assembled using SPAdes v3.13.1 (Prjibelski et al. 2020) prior to be used in our study.

precise locations on a reference genome and uses the positions to concatenate the overlapping k-mers into long-mers (Denancé et al. 2019). A k-mer size of 22 was used to search for specific sequences using the dataset with 118 genome sequences (Supplementary Table S1). Specific long-mers were searched in clusters 2ba-A, 2ba-B, 2ba-C, 2a-D, and 2a-E; in larger groups 2ba-ABC and 2a-DE encompassing the clusters of VCZ strains; and in clades 2b and 2d (Supplementary Table S1). Different groups of clade-2b strains were tested to isolate specific long-mers, including and excluding the pathovar *pisi*. The precise composition of in- and out-groups per marker is given in Supplementary Table S1. Depending on the target group, the size of long-mers varied from 22 to several thousand bp. To design PCR and qPCR tests, only long-mers with a size $\geq$ 70 bp were considered.

Primer3 2.3.4 was used to select primers in specific long-mers for all classical PCR assays with parameters set up for an optimal primer size of 20 bp, a product length of 150 to 350 bp, and a primer melting temperature (Tm) of 60 $\pm$ 3°C.

Development of a multiplex qPCR test to identify VCZ strains

A multiplex TaqMan qPCR test combining primers and probes targeting *sylC*, the two genes encoding the pathogenicity factors AvrRpt2 and HopZ5, and long-mers specific to clade-2d strains was developed to detect all VCZ strains and gain information on the strains present in the samples. Target sequences (*avrRpt2*, *hopZ5*, and long-mers specific to clade 2d, see above) were extracted from genome sequences by BLAST using the bioinformatic tool Extract_Genes_Genomes (https://iris.angers.inra.fr/galaxypub-cfbp) and aligned using Geneious R9. Primers and probe combinations for qPCR assays were designed within these specific regions using Geneious 9.1.8 (Koressaar and Remm 2007). The parameters were set up with an optimal primer size of 20 bp, an optimal product size of 80 bp, and Tm of 60 and 70°C for the primers and probe, respectively. Primer3 release 2.4.0 was used to design the qPCR assay for clade-2b strains. Dimer- and hairpin-forming primers and probes were discarded.

In silico evaluation of the specificity of the PCR and qPCR tests

Specificity of primers and associated probes was tested in silico using PrimerSearch (Val Curwen, Human Genome Mapping Project, Cambridge, U.K.), first on our collection of 121 genome sequences, and second on the 194,438 Whole Genome Shotgun (WGS) sequences of Bacteria and Archaea available in the NCBI database (March 2019 release). When necessary, the phylogenetic position of strains was checked using KISonWGS (https://iris.angers.inra.fr/galaxypub-cfbp). KISonWGS identifies the genome sequences in the NCBI database that are the closest neighbors of the genome sequences in the input query file. For multiplex the qPCR assays, sets of primers and probes, and combinations of primers and probes were tested on the genome sequences of $P99_{2ba\text{-}A}$, $P12857_{2ba\text{-}B}$, $ZUM3584_{2ba\text{-}C}$, $P77_{2a\text{-}D}$, $P66_{2a\text{-}E}$, $13\text{-}139B_{2a\text{-}F}$, $P108_{2b}$, and $P135_{2d}$ using AmplifX version 2.0.7 (Nicolas Jullien, Aix-Marseille Univ, CNRS, INP, Inst Neurophysiopathol, Marseille, France; https://inp.univ-amu.fr/en/amplifx-manage-test-and-design-your-primers-for-pcr) to check for the absence of dimers and cross-amplifications.

PCR and qPCR assays, specificity, and sensitivity

Classical PCR assays were performed in a 20-µl volume containing 200 nM of each primer, 200 nM of dNTP, 0.08 µl of GoTaq2, 4 µl of the 5× corresponding buffer (Promega), and 5 µl of boiled bacterial suspension at 1×10^7 CFU ml^{-1}. Amplification reactions were run in a thermocycler for 5 min at 95°C; 35 cycles of 30 s at 94°C, 30 s at 57°C, and 1 min at 72°C; and 3 min at 72°C. Five µl of sample were used for gel electrophoresis for 20 min at 100 V in a 1% agarose gel stained with ethidium bromide and visualized on a UV transilluminator.

All qPCR assays were performed in a Bio-Rad CFX96 Touch thermocycler and analyzed with Bio-Rad CFX Manager 3.1 software. Each individual and multiplexed TaqMan reaction was performed in a final volume of 10 µl containing 5 µl of Sso Advanced Universal Probes Supermix (Bio-Rad, Marne-la-Coquette, France), 600 nM of each primer, 200 nM of probe, and 1 µl of boiled, calibrated bacterial suspension. The amplification program was 3 min at 95°C, followed by 40 cycles of 15 s at 95°C and 30 s at 60°C. The qPCR assay for *sylC* was also performed in a final volume of 20 µl containing 10 µl of MasterMix buffer (Eurogentec, Seraing, Belgium), 900 nM of each primer, 250 nM of probe, and 5 µl of boiled, calibrated bacterial suspension. The amplification program was 10 min at 95°C followed by 40 cycles of 15 s at 95°C and 15 s at 60°C. For each reaction, the standard curve, the efficiency of the qPCR assay ($E = 10^{(-1/\text{slope})}$), and the correlation coefficient, R^2, were determined using calibrated DNA solutions (1 µg ml^{-1} to 1 pg ml^{-1}) of strains $P99_{2ba\text{-}A}$ for the qPCR assays for *sylC* and *avrRpt2* (primers and probes 4Ba and Rpt2, respectively), $P66_{2a\text{-}E}$ for the qPCR assay for *hopZ5* (primers and probe Z5), and $P108_{2b}$ and $P129_{2d}$ for the qPCR assays for strains in clades 2b* and 2d (primers and probes 2b* and 2d, respectively).

The specificity of the primers and probes was tested experimentally with the collection of target and nontarget bacterial strains (Table 1). Specificity was checked in the simplex and multiplex assays at least twice in duplicate samples using boiled bacterial suspensions ($OD_{650\text{ nm}} = 0.1$).

The sensitivity of the multiplex qPCR assays was evaluated with spiked seed samples. Batches of 500 seeds were soaked in buffer to produce the seed extract. Boiled bacterial suspensions of strains $P99_{2ba\text{-}A}$, $P66_{2a\text{-}E}$, and $P135_{2d}$ were inoculated at final concentrations ranging from 1×10^7 to 10 CFU ml^{-1} in 1 ml of seed extracts. For each strain, two independent experiments were done, including triplicate samples. DNA was extracted from 1-ml samples, and each DNA extract eluted in a 30-µl volume. Using this process, samples were concentrated 33-fold. Samples (1 µl per reaction) were amplified in triplicate with the multiplex qPCR tests. The limit of detection (LOD) was the lowest concentration of a seed extract with a positive signal in at least two of the triplicate samples.

Results

The seven-gene MLSA scheme revealed five clusters of VCZ strains in our collection

The analysis of our strain collection using the seven-gene MLSA scheme revealed that 29 strains belonged to cluster 2ba-A, 7 to cluster 2ba-B, 2 to cluster 2ba-C, 9 to cluster 2a-D, and 6 to cluster 2a-E (Fig. 1). VCZ strains were distributed in five clusters that all included strains responsible for other diseases on cucurbits (Fig. 1). In this study, no VCZ strain belonged to cluster 2a-F. The strains in clade 2b and 2d that were analyzed in the MLSA were genetically diverse and did not form any clusters.

Genomic sequence analysis of clade-2d strains revealed different T3E repertoires from VCZ strains

Strains $P129_{2d}$ and $P135_{2d}$ isolated from zucchini seed lots belonged to clade 2d, as revealed by the MLSA (Fig. 1) and confirmed by genome sequence analysis (Fig. 2). The distribution of T3E-encoding genes revealed diversified repertoires among the different phylogroups and clades but some common features among VCZ strains, such as eight conserved T3Es (Fig. 3; Supplementary Table S3). Genome sequences of strains $P129_{2d}$ and

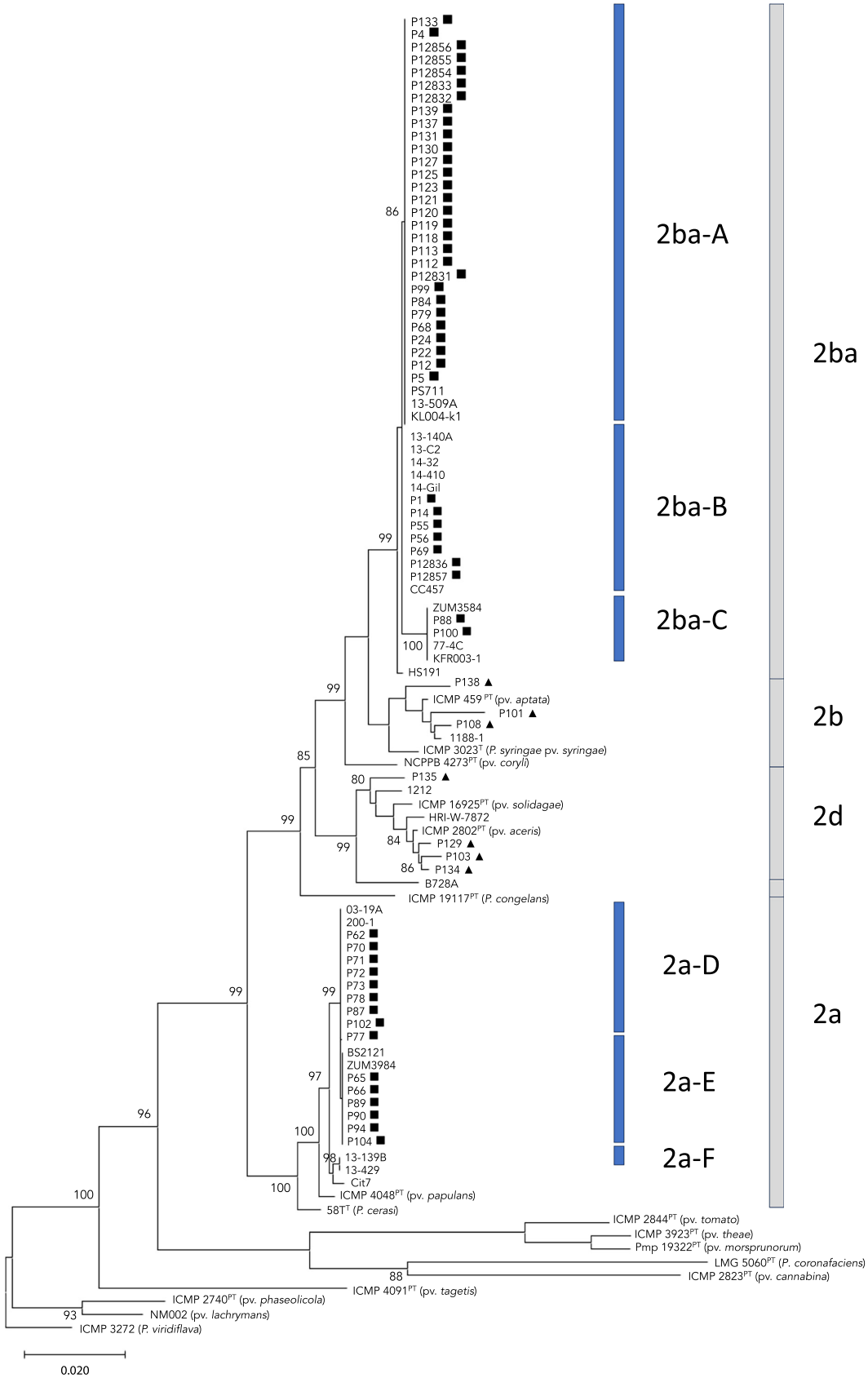

FIGURE 1

Phylogenetic relationships of strains of *Pseudomonas syringae* isolated from seed lots from 2002 to 2014. The maximum likelihood (ML) tree was based on concatenated partial sequences of *gapA*, *gltA*, *gyrB*, *rpoD*, *Psyr3208*, *Psyr3420*, and *Psyr4880* (3,405 bp). Bootstrap scores in percentage (1,000 replicates) are displayed at each node. Clades are indicated based on several studies (Berge et al. 2014; Bull and Koike 2015; Newberry et al. 2019), and clusters according to Lacault et al. (2020). Black symbols indicate strains from the collection used for this study: squares represent strains responsible for vein clearing of zucchini, and triangles represent clade-2b and -2d strains isolated from zucchini seeds.

P135$_{2d}$ contained 11 and 12 T3E encoding genes, respectively (Fig. 3), of which *avrE1*, *hopAA1*, *hopAE1*, *hopAG1*, *hopAH1*, and *hopI1* were present in all VCZ strains. The T3E gene *hopH1* was shared with more than 95% of VCZ strains, *hopBC1* with all clade-2ba strains, and *hopM1* with all VCZ strains except those of clade 2a-D. The T3E genes *hopA2* and *hopAY1* were present in both strains whereas *hopBE1* was only present in P135$_{2d}$. These clade-2d strains lacked *hopAZ1*, *hopBA1*, and *hopA1*, which were present in VCZ strains of clades 2a and 2ba (Lacault et al. 2020). The T3E genes *avrRpt2* and *hopZ5*, one of which was present in the genome sequences of all VCZ strains, were not present in P129$_{2d}$ and P135$_{2d}$ genome sequences. Similarly, the singleton P108$_{2b}$ had a small T3E repertoire and lacked *hopA1*, *avrRpt2*, and *hopZ5* (Lacault et al. 2020). This genome mining also revealed the presence of *sylC* in the genomes of all VCZ strains, and in some strains belonging to clades 2a and 2b (Fig. 2).

The k-mer–based comparison of genomic sequences enabled design of PCR and qPCR tests for all targeted strain groups except clade 2b

Specific sequences of each of the clusters 2ba-A to 2a-E, of larger groups 2ba-ABC (including clusters 2ba-A, -B, and -C) and 2a-DE (including clusters 2a-D and -E), and of clades 2b and 2d, were identified in the genome sequences using SkIf_with_DSK (*data not shown*). Their number and size varied depending on the targeted group. Among the 82 long-mers specific to cluster 2ba-A, the 45 specific to cluster 2ba-B, the 147 specific to cluster 2ba-C, the 124 specific to cluster 2a-D, the 213 specific to cluster 2a-E, the 239 specific to group 2ba-ABC, the 377 specific to group 2a-DE, and the 2,554 specific to clade 2d, the first sequence of each targeted group was generally used as a template sequence for PCR primers or qPCR primers and probe design (Table 3). Concerning clade 2b, only two specific

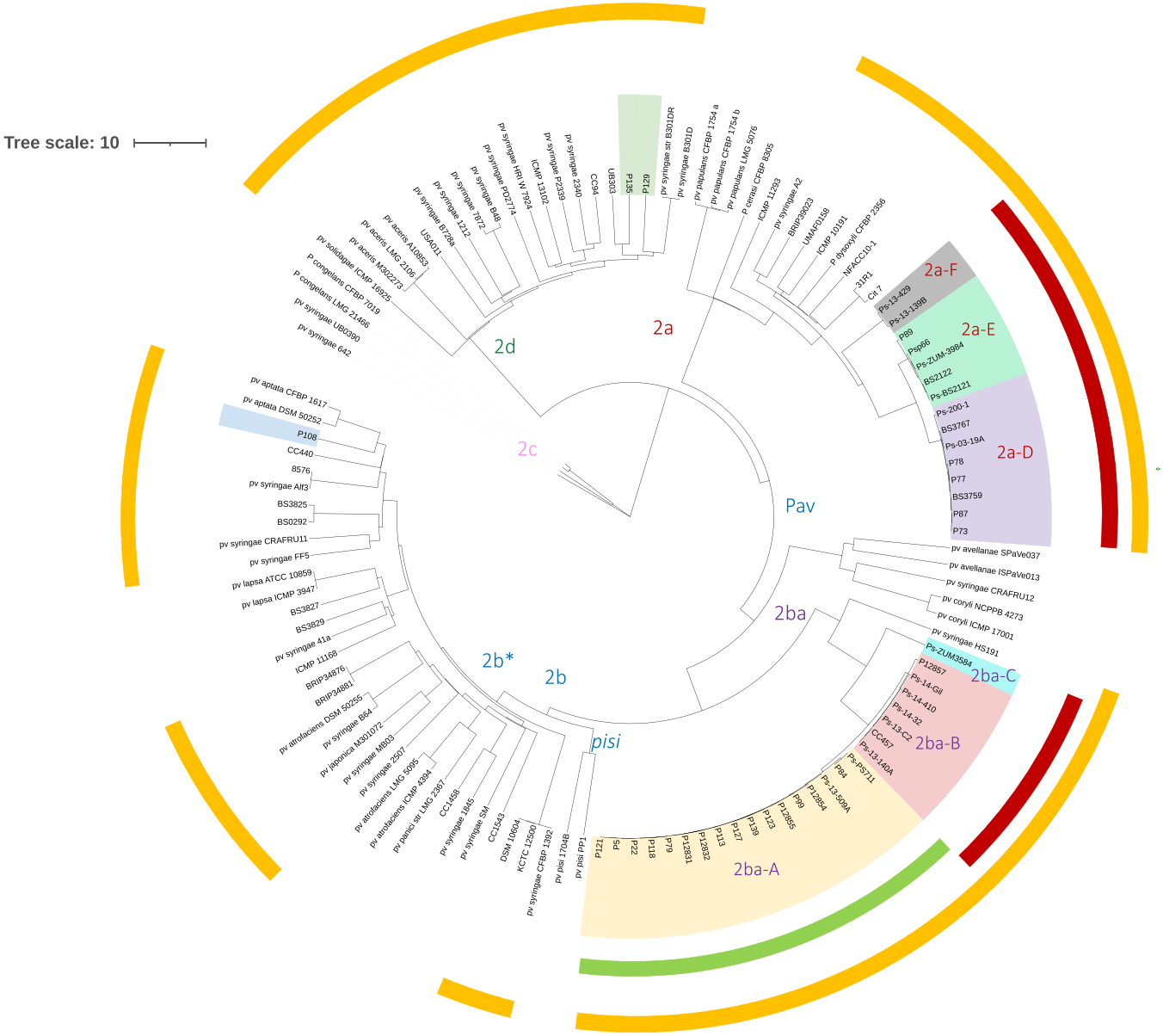

FIGURE 2

Phylogenetic groups chosen for PCR assay design. The dendrogram was built on distances calculated from shared k-mer matrices obtained with KIS on a dataset of genomic sequences of 118 strains of *Pseudomonas syringae* phylogroup 2. Clades are indicated based on several studies (Berge et al. 2014; Bull and Koike 2015; Newberry et al. 2019), and clusters according to Lacault et al. (2020). The presence of *avrRpt2*, *hopZ5*, and *sylC* is indicated by green, red, and yellow arcs of the outside circles, respectively.

long-mers enabled qPCR primers and probe design, identified in the node 2b*, excluding pathovar *pisi* (Fig. 2; Table 3). No specific primers could be designed for sequences in the clade-2b node, even when adding clades 2ba or Pav to the in-group, because (i) of the lack of specific long-mers, (ii) long-mers were not appropriate to design a test, or (iii) the test was not inclusive of clade-2b strains isolated from zucchini (*data not shown*).

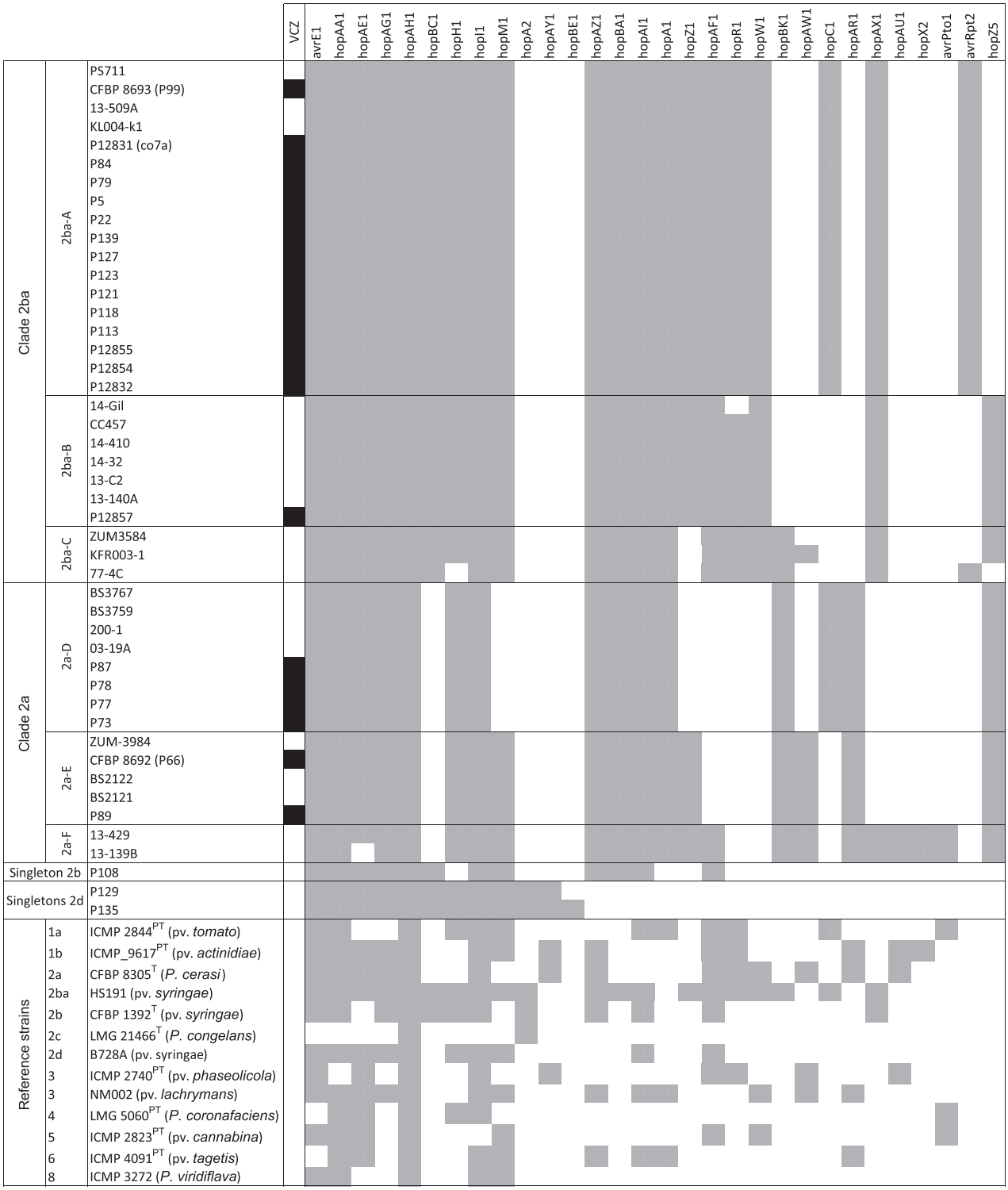

FIGURE 3

Type three effector (T3E) genes in strains isolated from zucchini and their distribution among representative strains of the *Pseudomonas syringae* species complex. A black square indicates a strain responsible for vein clearing of zucchini (VCZ). T3Es were searched using tBLASTN with the genome sequences. A gray square indicates the presence of a sequence with an identity ≥ 80% on at least 80% of the protein length. Clades are indicated based on several studies (Berge et al. 2014; Bull and Koike 2015; Newberry et al. 2019), and clusters according to Lacault et al. (2020).

PCR and qPCR tests detected in silico and experimentally all targeted genome sequences or strains

Inclusivity is defined as the percentage of target microbial genomes or strains that gave the expected positive result, and exclusivity as the percentage of nontarget microbial genomes or strains that gave the expected negative result (European and Mediterranean Plant Protection Organization 2019). The PCR tools developed to determine the position of strains in the different clusters and clades were 100% inclusive experimentally. Inclusivity was also excellent in silico, as only one exception was identified, for one strain in clade 2b* (Table 4). A limited number of cross-reactions were observed, resulting in an exclusivity of 99.96 to 100%, depending on the test (Supplementary Table S2; Tables 1 and 4). Several of these cross-reactions involved strains with in-between genotypes, revealing a greater diversity than expected within clade 2ba. For the test targeting the cluster 2ba-A, strain 1188-1 (RBPG00000000) isolated from zucchini, classified in clade 2b on the basis of the whole genome sequence, was positive in silico with the 2ba-A test (Table 4). This genome sequence was also positive in silico for other PCR tests targeting cluster 2ba-A, such as 2ba-ABC, 4Ba, and Rpt2 (targeting *avrRpt2*) PCR tests, and matched the primers and probe for clade 2b (Supplementary Table S2). Similarly, positive reactions with 2ba-A and Rpt2 primers were observed in silico and experimentally for three strains belonging to clusters 2ba-C2 and 2ba-B (Supplementary Table S2; Table 1). For the PCR test dedicated to cluster 2ba-B strains, three cross-reactions were observed in silico for *P. coronafaciens* genomes (PG4; Table 4), which were confirmed experimentally by a positive signal for the type strain of *P. coronafaciens* (Table 1). For cluster 2ba-C, primers annealed with four mismatches in silico with a strain of clade 2a and gave a band of the wrong size experimentally for one clade-2b strain. Two genome sequences of *P. coronafaciens* perfectly matched with 2a-D primers. Primers for group 2ba-ABC annealed in silico with mismatches for two strains, including strain 1188-1. Primers for group 2a-DE annealed with the two cluster 2a-F genome sequences, with three mismatches. Concerning the 2b* PCR test, the clade-2b genome sequence of strain ICMP 4917 (RBUE00000000) did not match in silico. Numerous in silico imperfect matches were obtained for the primers and probe targeting clade 2b* of nontarget genomes but were not confirmed experimentally with representative strains (Supplementary Table S2). The genome sequence of strain ICMP 4996 (RBSD00000000), named *P. coronafaciens* pv. *striafaciens* in the NCBI database, was classified in clade 2b according to the KISonWGS assignment and to Dillon et al. (2019b), but not in PG4 as expected based on the taxonomic assignment. This genomic sequence perfectly matched the 2b* primers and probe. This result was confirmed experimentally with the synonymous strain CFBP 1750 that gave a positive signal with this test. This particular case was considered a problem of misclassification and not a true cross-reaction. The primers and probe dedicated to clade-2d strains gave one cross-reaction in silico with the clade-2b strain 3023 (MLFD00000000). For cross-reactions that did not involve strains with intermediate genotypes, a preliminary detection with the *sylC*-based PCR assay helped rule them out (Supplementary Table S2).

In silico and experimental testing of the 4Ba qPCR test confirmed that this test was 100% inclusive of VCZ strains but also detected some other PG2 strains despite mismatches in

TABLE 3

PCR and qPCR primers and probes designed to detect and identify strains belonging to the species complex *Pseudomonas syringae*, isolated from zucchini seeds

Target group	Primers-probe name[a]	Sequence of primers and probes (5′–3′)	Amplicon size (bp)	Annealing temperature (°C)	Target sequence (protein ID)[b]
2ba-A	2ba-A-F	GACCAACCTGCCCATTTGAG	163	57	PspCFBP8693_11730
	2ba-A-R	GGCCCTTAACGTGAGAGAGT		57	(NAP44764.1)
2ba-B	2ba-B-F	GGCAGGATCTCGAACAGTTC	160	57	PscP12857CL_10745
	2ba-B-R	CCTTCTCGACCCACCAAGTC		57	(NAO47743.1)
2ba-C	2ba-C-F	GCTAACGCTTGATTCTCGGT	198	57	B1F68_00160
	2ba-C-R	GCCGTGTATGCCTAACCAGA		57	(RXU12171.1)
2a-D	2a-D-F	GATGGAGTGCAGACATCACG	353	57	PscP77CL_ RS08930
	2a-D-R	CCAAATTCCGTCAATTGGTT		57	(NAP49229.1)
2a-E	2a-E-F	CTCCACGAATGGGTTGAAGT	353	57	PscCFBP8692_13650
	2a-E-R	TACCCTTGTGAATGGCAACA		57	(NAP19626.1)
2ba-ABC	2ba-ABC-F	TCCGGAACAGTGACACGTAA	219	59	PspCFBP8693_18355
	2ba-ABC-R	TAGAACAAGCGCGCATTACC		59	(NAP46023.1)
2a-DE	2a-DE-F	GAGCGTTACGTCCTTGAAGC	350	57	PscCFBP8692_RS07980
	2a-DE-R	CGTCTGGAGCTTCTTGTTCC		57	(NAP18542.1)
2b*[c]	2b*-F	TCAAGACGCTGGAAAATGCA	129	58	PsP108CL_23575
	2b*-P	^{FAM}GCTGTTATTGCTGCTGCTCT^{BHQ1}		60	(NAO55204.1)
	2b*-R	AGGATCGGAATAATTGCAGTGG		59	
2d	2d-F	CGTGCTGCTCAAATCGCC	80	59.9	KBY08_08415
	2d-P	^{Cy5}CGCGACGGTGCTCGGTTCGC^{BHQ2}		69.1	(MCA5971736.1)
	2d-R	TCCTGTCAGCCATTGGGG		58.9	
Strains with *sylC*	4Ba-F2	CTTTGAGTACGCCCAAGAGC	152	56	*sylC*
	4Ba-P	^{FAM}TCGGAGCGCATCATGCAGAT^{BHQ1}		59	CFBP8693_RS03260
	4Ba-R3	GGAGGCAGTGCTGACGTA		56	(NAP43145.1)
Strains with *hopZ5*	Z5-F	GCCTGGCTATCTTCAGATGC	79	58.5	*hopZ5*
	Z5-P	^{TxRed}ACGCTCAGGATAACGGCCGC^{BHQ2}		65.3	PscCFBP8692_RS27610
	Z5-R	TCCAAGTTGGATGACAGCCA		59.2	(NAP22298.1)
Strains with *avrRpt2*	Rpt2-F	CTGCGCAAGAATGGTTGGC	80	60.4	*avrRpt2*
	Rpt2-P	^{HEX}GCCTCGCCTAGGGCTGCCGG^{BHQ1}		69.9	PspCFBP8693_RS11460
	Rpt2-R	GGCCCTCCCTTCCCTCAT		60.4	(NAP44711.1)

[a] Primers forward (F) and reverse (R) and probe (P) for PCR and TaqMan qPCR.
[b] Gene ID (Protein ID) in NCBI.
[c] 2b* corresponded to part of clade 2b that did not include strains belonging to clade 2ba or to pathovars *pisi*, *avellanae*, and *coryli* (Fig. 2).

the primers (Table 1; Supplementary Table S2). The primers and probe targeting the gene *avrRpt2* annealed to the genome sequences of all strains in clusters 2ba-A and 2ba-C2, strain 1188-1_{2b}, and some strains of pathovars *tomato*, *spinaceae*, *maculicola*, and *tagetis*. These in silico results were confirmed experimentally with the detection of one strain of pathovar *tomato* and another one of pathovar *spinaceae*. The Z5 primers and probes targeting *hopZ5* detected in silico and experimentally strains in clusters 2ba-B and -C and 2a-D, -E, and -F, and strains of the pathovar *actinidiae*. This later pathovar is known to have *hopZ5*. Imperfect matches were found in silico for pathovars *berberidis*, *delphinii*, and *phaseolicola*, but experiments with representative strains were negative. All strains of those cross-reacting pathovars were negative with the 4Ba qPCR test.

The multiplex qPCR test identified VCZ strains and gave insight into their pathogenicity

All targeted sequences, when present in a genome sequence, were found in single copy. Individually, the three PCR tests targeting genes encoding the pathogenicity factors SylC, AvrRpt2, and HopZ5 were not specific for VCZ strains (Supplementary Table S2). However, VCZ strains from the main clusters are characterized by the simultaneous presence of *sylC* and either *avrRpt2* or *hopZ5* (Fig. 2; Table 1), as *avrRpt2* characterizes 2ba strains that have a narrow host range on cucurbits whereas *hopZ5* is present in strains with a wider host range (Djitro et al. 2022a; Lacault et al. 2020). Clade-2b* and -2d strains were also isolated from zucchini seeds and gave a positive signal with the 4Ba PCR test targeting *sylC*, justifying the need to include their corresponding qPCR tests in the multiplex assay. However, the attempt to add the clade-2b* qPCR test in the multiplex was abandoned due to qPCR quantitative parameters that were incompatible with other tests (*data not shown*).

The specificity of each qPCR test was not affected by multiplexing. In silico simulation of the multiplex qPCR test with AmplifX did not show any cross reactions between the different pairs of primers and the probes. The efficiencies of the qPCR tests were evaluated in simplex and multiplex using serial dilutions of DNA extracted from each target strain. Parameters such as PCR efficiency and R^2 of regression standard curves were correct in both the simplex and multiplex assays, except for the 4Ba test

TABLE 4

In silico and experimental specificity of PCR and qPCR assays on the whole NCBI database, and on a collection of *Pseudomonas syringae* strains responsible for vein clearing of zucchini (VCZ) and other bacterial strains

		Number of genome sequences with in silico annealings		Number of strains with positive signals in experimental PCR assays	
PCR or qPCR test	Targeted group	Target (inclusivity)[a]	Nontarget (exclusivity)[b]	Target (inclusivity)[a]	Nontarget (exclusivity)[b]
2ba-A	2ba-A	18 (100%)	2[c] (99.999%)	30 (100%)	2[d] (97.56%)
2ba-B	2ba-B	8 (100%)	13[e] with 0 to 8 mismatches (99.993%)	7 (100%)	1[f] (99.05%)
2ba-C	2ba-C	3 (100%)	7[g] with 4 to 8 mismatches (99.996%)	2 (100%)	0 (100%)
2a-D	2a-D	8 (100%)	2[h] (99.999%)	9 (100%)	0 (100%)
2a-E	2a-E	5 (100%)	0 (100%)	6 (100%)	0 (100%)
2ba-ABC	2ba-A, 2ab-B, 2ba-C	26 (100%)	2[i] with 3 to 5 mismatches (99.999%)	39 (100%)	0 (100%)
2a-DE	2a-D, 2a-E	13 (100%)	2[j] with 3 mismatches (99.999%)	15 (100%)	0 (100%)
2b*[k]	2b*[k]	60[l] (98.36%)	0 (100%)	9 (100%)	0 (100%)
2d	2d	64 (100%)	1[m] (99.999%)	18 (100%)	0 (100%)
4Ba[n] + Rpt2	2ba-A, 2ba-C2	19 (100%)	1[c] (99.999%)	30 (100%)	2[d] (97.56%)
4Ba + Z5	2ba-B, 2ba-C1, 2a-D, 2a-E, 2a-F	25 (100%)	0 (100%)	22 (100%)	0 (100%)
4Ba + 2b*[k]	2b*[k] with *sylC*	30[l] (96.77%)	0 (100%)	8 (100%)	0 (100%)
4Ba + 2d	2d with *sylC*	64 (100%)	1[m] (99.999%)	18 (100%)	0 (100%)

[a] Inclusivity (European and Mediterranean Plant Protection Organization 2019) was defined as the percentage of target microbial genomes or strains that gave the expected positive result.

[b] Exclusivity (European and Mediterranean Plant Protection Organization 2019) was defined as the percentage of nontarget microbial genomes or strains that gave the expected negative result. The number of nontargets was 194,450 genomes of the 2019-NCBI release minus the targets (Supplementary Table S2), and the 112 strains of our collection for experimental test minus the targets as indicated (Table 1).

[c] Unexpected perfect match of primers was found on genomic sequences of strain 1188-1 (RBPG00000000) belonging to clade 2b and of strain 77-4C (SRR21657462) belonging to cluster 2ba-C2.

[d] Unexpected positive result for strains $P55_{2ba-B}$ and $P56_{2ba-B}$ classified in cluster 2ba-B by multilocus sequence analysis (Fig. 1; Table 1).

[e] Unexpected perfect match of primers was found with genome sequences of three strains of *P. coronafaciens*, including the type strain LMG 5060^T (RBUJ00000000, RBPK00000000, and JSED00000000), and unexpected match with seven to eight mismatches with 10 genome sequences belonging to various genera delimiting a fragment with a size close to the expected target.

[f] Unexpected positive result for the type strain of *P. coronafaciens*, strain CFBP 2216^T (homolog of LMG 5060^T).

[g] Unexpected match with four mismatches was found with the genome sequence of strain G733 (RBOI00000000) belonging to clade 2a and unexpected match with eight mismatches with six genome sequences belonging to various genera delimiting a fragment with a size close to the expected target.

[h] Unexpected perfect match of primers with genome sequences of two strains of *P. coronafaciens* pv. *oryzae* belonging to phylogroup 4 (PG4) (RBSZ00000000, RBSY00000000).

[i] Unexpected matches with three and five mismatches were found with genome sequences of strain 1188-1 (RBPG00000000) belonging to clade 2b and an unclassified strain of *Pseudomonas* sp. (MPIP00000000).

[j] Unexpected match with three mismatches was found with genome sequences of two strains belonging to cluster 2a-F, strains 13-429 and 13-139B (MVAY00000000 and MVAT00000000).

[k] 2b* is part of clade 2b without strains belonging to pathovar *pisi* (Fig. 2).

[l] Primers and probe did not match one genome sequence of strain ICMP 4917 (RBNV00000000) that belongs to the target group (part of clade 2b as defined above).

[m] Unexpected perfect match of primers with genome sequence of strain 3023 (MLFD00000000) that belongs to clade 2b.

[n] These primers and probe perfectly matched with 66 genome sequences and with 95 genome sequences with 5 mismatches in the primers. They gave a positive signal experimentally for 81 PG2 strains of the collection.

that had better efficiency in the conditions of the multiplex than the simplex qPCR assays (Table 5). Specificity of each qPCR test was identical in multiplex and simplex (Table 1).

When using calibrated DNA solutions, the LOD of each qPCR test in the multiplex assay was 100 pg ml^{-1} of DNA (0.1 pg per reaction, corresponding to 16 genome copies), but was 10 times better in the simplex qPCR assays for 4Ba, Rpt2, and Z5 (Table 5). When determined using boiled bacterial suspensions, the LODs for the multiplex assay were 2.04×10^4, 2.46×10^4, 4×10^4, and 5.80×10^3 CFU ml^{-1} (corresponding to 20, 25, 40, and 6 CFUs per reaction) for the 4Ba-, Rpt2-, Z5-, and 2d-qPCR assays, respectively (Supplementary Table S4). Similar results were obtained in the second experiment. The apparently better LOD obtained for the 2d-qPCR test with bacterial suspensions may be due to the higher efficiency of the 2d-qPCR assay compared to the tests run with the other primer pairs. Differences were even higher for bacterial suspensions than for DNA solutions (Table 5; Supplementary Table S4).

Multiplex qPCR assay in combination with a DNA extraction protocol that can be automated enabled detection of VCZ strains in seed lots

The LODs of the spiked seed extracts corresponded to 2.46×10^3, 2.46×10^3, 3.22×10^3, and 5.80×10^2 CFU ml^{-1} for the qPCR assays for 4Ba, Rpt2, Z5, and 2d, respectively (Table 6). The parameters of the standard curve for DNA extracted from spiked seed extracts (Table 6) were not different from those obtained using pure bacterial suspensions (Supplementary Table S5). Except for strain $P135_{2d,}$ these LODs were 10 times lower than the LODs obtained for calibrated bacterial suspensions (Table 6; Supplementary Table S4) due to the increase in DNA concentration during extraction and the absence of PCR inhibitors

TABLE 5

Efficiencies of simplex and multiplex qPCR tests on calibrated DNA extracts of strains of *Pseudomonas syringae* representative of the target groups

qPCR test (strain)	DNA concentration	Number of genome copies per milliliter[a]	Mean Cq value (SEM)[b]	
			Simplex	Tetraplex
4Ba ($P99_{2ba-A}$)	1 µg ml^{-1}	1.65×10^8	22.43 (0.29)	19.43 (0.09)
	100 ng ml^{-1}	1.65×10^7	26.05 (0.13)	22.85 (0.09)
	10 ng ml^{-1}	1.65×10^6	29.13 (0.07)	26.36 (0.12)
	1 ng ml^{-1}	1.65×10^5	32.69 (0.39)	29.37 (0.05)
	100 pg ml^{-1}[c]	1.65×10^4	37.98 (1.93)	33.26 (0.46)
	10 pg ml^{-1}[d]	1.65×10^3	38.95 (0.26)	NA
	1 pg ml^{-1}	1.65×10^2	NA	NA
Efficiency			82%	94.9%
R^2			0.963	0.989
Slope			−3.84	−3.451
Rpt2 ($P99_{2ba-A}$)	1 µg ml^{-1}	1.65×10^8	21.55 (0.15)	19.77 (0.08)
	100 ng ml^{-1}	1.65×10^7	25.11 (0.12)	23.19 (0.08)
	10 ng ml^{-1}	1.65×10^6	28.49 (0.12)	26.64 (0.10)
	1 ng ml^{-1}	1.65×10^5	32.47 (0.17)	30.01 (0.16)
	100 pg ml^{-1}[c]	1.65×10^4	35.47 (0.21)	33.93 (0.35)
	10 pg ml^{-1}[d]	1.65×10^3	39.05 (0.31)	NA
	1 pg ml^{-1}	1.65×10^2	NA	NA
Efficiency			92.9%	92.3%
R^2			0.997	0.99
Slope			−3.505	−3.521
Z5 ($P66_{2a-E}$)	1 µg ml^{-1}	1.62×10^8	21.18 (0.08)	18.34 (0.12)
	100 ng ml^{-1}	1.62×10^7	24.20 (0.05)	22.25 (0.27)
	10 ng ml^{-1}	1.62×10^6	27.54 (0.12)	25.36 (0.12)
	1 ng ml^{-1}	1.62×10^5	30.95 (0.04)	28.79 (0.17)
	100 pg ml^{-1}[c]	1.62×10^4	34.22 (0.24)	33.11 (0.01)
	10 pg ml^{-1}[d]	1.62×10^3	36.68 (0.67)	*35.91*
	1 pg ml^{-1}	1.62×10^2	NA	NA
Efficiency			95.8%	90.2%
R^2			0.991	0.995
Slope			−3.426	−3.582
2d ($P129_{2d}$)	1 µg ml^{-1}	1.61×10^8	26.27 (0.09)	19.99 (0.07)
	100 ng ml^{-1}	1.61×10^7	28.15 (0.41)	24.25 (0.16)
	10 ng ml^{-1}	1.61×10^6	32.03 (0.28)	27.66 (0.10)
	1 ng ml^{-1}	1.61×10^5	35.91 (0.61)	30.49 (0.27)
	100 pg ml^{-1}[c, d]	1.61×10^4	38.08 (1.01)	33.36 (0.29)
	10 pg ml^{-1}	1.61×10^3	*37.83*	*37.09*
	1 pg ml^{-1}	1.61×10^2	NA	NA
Efficiency			99.8%	101%
R^2			0.994	0.988
Slope			−3.326	−3.299

[a] The theoretical number of genome copies in DNA solutions was calculated with 1 pg corresponding to 978 Mb and genome sizes to 5.908, 6.185, and 6.044 Mb for $P99_{2ba-A}$, $P66_{2a-E}$ and $P129_{2d}$, respectively.

[b] NA = Not Amplified; DNAs were tested in triplicate with 1 µl per reaction; values in italics correspond to a unique Cq value; mean Cq values were calculated on at least two positive values.

[c] LOD of the qPCR test used in tetraplex qPCR assay.

[d] LOD for the qPCR test used in simplex qPCR assay.

after DNA extraction. Thus, the protocol used to extract bacterial DNA seemed to be adapted to discard PCR inhibitors from seed and recover bacterial DNA with a quality compatible with qPCR assays.

Based on the number of seeds per subsample, the number of subsamples, and the number of positive subsamples (Table 7), percentages of contaminated seeds in seed lots I to IV were estimated at 1.24, 1.24, 0.9, and 2.4%, respectively, according to Maury's formula for subsamples of the same size (Maury et al. 1986) and to Swaroop's tables for subsamples with decreasing sizes (Swaroop 1951). Seed lots I and II were contaminated at the same rate (1.24%). For each lot, bacterial concentrations were similar in the same subsamples for the *hopZ5* and *sylC* PCR tests. This strongly suggested infection of these two seed lots with a VCZ strain with a wide host range. Seed lot III was contaminated at 0.9%, probably with a clade-2d strain because population sizes estimated with both tests (2d and *sylC* PCR tests) were congruent in the subsamples. Seed lot IV corresponded to a seed lot treated with a disinfectant that had infection by VCZ strains that were resistant to disinfection. Because the number of genome copies estimated for *sylC* in the subsamples corresponded the sum of the genome copies estimated for *avrRpt2* and *hopZ5* markers, i.e., the sum of 0.91×10^6 CFU per 500 seeds (number of genomes having the *avrRpt2* marker) and 0.2×10^6 CFU per 500 seeds (number of genomes having the *hopZ5* marker) was 1.11×10^6 CFU, which is close to the 1.17×10^6 CFU per 500 seeds estimated for the marker *sylC*. This lot was probably infected with two strains, one that had a narrow cucurbit host range and was more abundant and a second strain that had a wider host range.

Discussion

Strains of *Pseudomonas syringae* are ubiquitous plant epiphytes and some are plant pathogens. They are diverse and, according to the weapons they dispose, such as T3SS and effectors, toxins, exopolysaccharides, cell wall-degrading enzymes, and plant hormones (Xin et al. 2018), they do not all have the same host range and aggressiveness. Depending on the diseases they cause, they do not represent the same threat for a crop and for the surrounding cultivated plants. The aim of the study was to provide tools to characterize, identify, detect, and quantify strains belonging to the *P. syringae* species complex that affect zucchini seed production. (i) PCR tests were designed to extend the previously proposed MLSA scheme to include a total of seven genes to characterize isolates from epidemics and establish their

TABLE 6

Sensitivity of the multiplex qPCR test on seed extracts spiked with calibrated bacterial suspensions of strains of *Pseudomonas syringae* representative of the target groups

	First experiment			Second experiment		
		Mean Cq (SEM)			Mean Cq (SEM)	
Strain	Concentration (CFU ml^{-1})[a]	4Ba	Rpt2	Concentration (CFU ml^{-1})[a]	4Ba	Rpt2
$P99_{2ba\text{-}A}$	2.46×10^7	18.16 (0.10)	18.11 (0.06)	3.28×10^7	18.18 (0.15)	17.91 (0.23)
	2.46×10^6	21.94 (0.21)	21.98 (0.17)	3.28×10^6	21.48 (0.24)	21.18 (0.18)
	2.46×10^5	25.53 (0.12)	25.74 (0.09)	3.28×10^5	25.19 (0.55)	25.02 (0.31)
	2.46×10^4	28.01 (0.01)	28.02 (0.05)	3.28×10^4	26.67 (0.17)	27.12 (0.04)
	2.46×10^3	31.75 (0.14)	31.59 (0.36)	3.28×10^3	32.5 (0.34)	31.36 (0.50)
	2.46×10^2	NA[b]	NA	3.28×10^2	NA	NA
	2.46×10^1	NA	NA	3.28×10^1	NA	NA
	0 to 3	NA	NA	0 to 4	NA	NA
Efficiency[c]		99.9%	100.9%		100.7%	102.9%
R^2		0.994	0.99		0.96	0.992
Slope		−3.325	−3.301		−3.306	−3.255
$P66_{2a\text{-}E}$	3.22×10^7	20.68 (0.03)	17.48 (0.06)	4×10^7	21.76 (0.12)	18.26 (0.06)
	3.22×10^6	24.81 (0.04)	21.37 (0.03)	4×10^6	25.61 (0.13)	22.11 (0.08)
	3.22×10^5	27.51 (0.12)	24.15 (0.04)	4×10^5	27.21 (0.13)	23.81 (0.14)
	3.22×10^4	31.31 (0.05)	26.56 (0.12)	4×10^4	33.92 (0.39)	27.82 (0.17)
	3.22×10^3	NA	30.89 (0.40)	4×10^3	NA	31.79 (0.64)
	3.22×10^2	NA	NA	4×10^2	NA	NA
	3.22×10^1	NA	NA	4×10^1	NA	NA
	0 to 4	NA	NA	0 to 4	NA	NA
Efficiency[c]		94.50%	105.2%		83.10%	110.8%
R^2		0.994	0.987		0.934	0.98
Slope		−3.461	−3.202		−3.808	−3.088
$P135_{2d}$	5.80×10^7	20.52 (0.12)	17.1 (0.13)	2.04×10^7	20.97 (0.16)	18.28 (0.14)
	5.80×10^6	23.25 (0.06)	19.93 (0.11)	2.04×10^6	23.78 (0.17)	20.99 (0.18)
	5.80×10^5	26.13 (0.01)	22.72 (0.28)	2.04×10^5	27.32 (0.14)	24.68 (0.32)
	5.80×10^4	30.30 (0.21)	26.79 (0.16)	2.04×10^4	31.78 (0.50)	27.51(0.13)
	5.80×10^3	NA	28.95 (0.08)	2.04×10^3	NA	31.09 (0.23)
	5.80×10^2	NA	33.28 (0.56)	2.04×10^2	NA	NA
	5.80×10^1	NA	NA	2.04×10^1	NA	NA
	0 to 4	NA	NA	0 to 3	NA	NA
Efficiency[c]		104.40%	107.6%		89.60%	104.7%
R^2		0.987	0.99		0.980	0.994
Slope		−3.221	−3.153		−3.598	−3.214

[a] Concentration in samples before DNA extraction. Seed extract samples were spiked with serial dilutions of each strain separately. DNAs were extracted from 1 ml of seed sample, eluted in 30 µl of buffer, and tested in triplicate with 1 µl per reaction.
[b] NA = not amplified.
[c] Standard curve and parameters of qPCR assays were determined in seed extract matrix.

phylogenetic relationships and taxonomy (Lacault et al. 2020). (ii) Based on a set of 118 genome sequences, PCR tests were developed to determine, without the need for sequencing, to which phylogenetic group the strains belong. These user-friendly tests are dedicated to working with large numbers of samples for epidemiological monitoring. (iii) In order to improve seed testing, the qPCR test based on *sylC* was multiplexed with other markers to identify the strains based on their pathogenicity to various Cucurbitaceae genera and species.

The new MLSA scheme based on seven genes enabled characterization of the whole collection of 50 VCZ strains and an additional 14 strains isolated from zucchini seeds and plants from seed production fields. Two VCZ strains were identified in the cluster 2ba-C, indicating that VCZ strains were present in five of the six clusters that were described previously for strains pathogenic on cucurbits (Lacault et al. 2020; Newberry et al. 2019). Clustering of strains in homogeneous lineages suggests they could be epidemic clones responsible for damage to zucchini seedlings (Lacault et al. 2020) and, for some, also on other cucurbits (Lacault et al. 2020; Newberry et al. 2019). MLSA revealed that strains belonging to clades 2b and 2d did not form clusters. The clade-2d strains caused symptoms on cucurbits but were less aggressive on zucchini plants than VCZ control strains. Strains $P129_{2d}$ and $P135_{2d}$ were discarded from a previous study (Lacault et al. 2020) because they were not considered VCZ strains. Indeed, they did not cause seedling deformation and stunting as did the VCZ control (Supplementary Fig. S1A). However, when tested, strain $P135_{2d}$ was able to cause water-soaked and necrotic lesions on leaves of various cucurbits after foliar spray inoculation, including on several squashes, melon, cucumber, and watermelon, as previously described (Lacault et al. 2020) (Supplementary Fig. S1B). Weak pathogenicity on squash was previously reported for clade-2d strains (Newberry et al. 2016). Genome analysis of clade-2d strains revealed small T3E repertoires similar to PG2 genome sequences in general (Dillon et al. 2019a). The strains had fewer T3Es in common with VCZ-cluster strains than the clade-2b strain $P108_{2b}$. Strains of clades 2b and 2d lacked *hopA1*, which is shared by all VCZ strains (Lacault et al. 2020) and was reported to be present in all strains responsible for epidemics on cucurbits in the United States (Newberry et al. 2019). Not falling into genetic clusters, clade-2b and -2d strains did not seem to be epidemic clones, even though they caused minor symptoms following inoculation. Increasing the number of these types of strains in the collection will be necessary to determine their potential role in causing diseases of zucchini.

To detect VCZ and related strains, we used genes encoding pathogenicity factors to improve the qPCR test based on *sylC*. This gene encodes a nonribosomal peptide synthetase module involved in the biosynthesis of syringolin A, a toxin secreted only by strains belonging to PG2 (Dillon et al. 2019b). The *sylC*

TABLE 7

Results of the analysis of naturally contaminated seed lots with the multiplex qPCR test

Seed lot[a]	Number of seeds per subsample[b]	Mean log genome copies per seed subsample (SEM)[c]			
		4Ba	Z5	Rpt2	2d
I	100	5.85 (0.05)	5.66 (0.03)	NA	NA
	100	5.34 (0.38)	5.63 (0.02)	NA	NA
	100	5.58 (0.33)	5.76 (0.01)	NA	NA
	100	5.85 (0.23)	5.77 (0.03)	NA	NA
	100	5.30 (0.41)	5.52 (0.05)	NA	NA
	100	NA[d]	NA	NA	NA
	100	NA	NA	NA	NA
II	100	5.82 (0.21)	5.83 (0.00)	NA	NA
	100	5.77 (0.39)	5.85 (0.01)	NA	NA
	100	5.99 (0.25)	5.99 (0.02)	NA	NA
	100	5.78 (0.18)	5.90 (0.04)	NA	NA
	100	5.58 (0.45)	5.81 (0.02)	NA	NA
	100	NA	NA	NA	NA
	100	NA	NA	NA	NA
III	500	5.57 (0.00)	NA	NA	5.62 (0.05)
	100	5.66 (0.55)	NA	NA	5.54 (0.27)
	100	5.19 (0.89)	NA	NA	6.11 (0.76)
	100	5.64 (0.00)	NA	NA	5.39 (0.25)
	100	NA	NA	NA	NA
	100	NA	NA	NA	NA
IV	500	6.07 (0.18)	4.32 (0.02)	5.96 (0.64)	NA
	100	5.17 (0.35)	2.79 (0.60)	4.88 (0.35)	NA
	100	5.16 (0.36)	2.57 (0.19)	4.77 (0.36)	NA
	100	5.50 (0.27)	3.52 (0.40)	5.07 (0.34)	NA
	100	5.01 (0.34)	3.50 (0.36)	4.54 (0.36)	NA
	100	5.48 (0.31)	3.56 (0.41)	5.34 (0.51)	NA
	10	NA	NA	NA	NA
	10	NA	NA	NA	NA
	10	NA	NA	NA	NA
	10	NA	NA	NA	NA
	10	NA	NA	NA	NA

[a] Seed lots I to III were produced in France in 2019 and IV in China in 2020.
[b] First, subsamples of 100 seeds were analyzed. If they were all negative or positive, bigger or smaller samples were analyzed in order to determine a percentage of contaminated seeds in the seed lot.
[c] Number of genome copies per sample of 100 or 500 seeds was the number genome copies in 1 µl of DNA used in the qPCR reaction multiplied by the total volume of extracted DNA and by the total volume of seed soaking. Samples were amplified in two different experiments, and two different equations were used to calculate the number of genome copies per reaction.
[d] NA = not amplified.

qPCR test (4Ba) detected all VCZ strains clustered in clades 2a and 2ba and also some other strains within PG2, in particular all clade-2d strains and a large proportion of clade-2b strains (Table 1; Supplementary Table S2). The qPCR test based on *sylC* was multiplexed with qPCR tests targeting *avrRpt2* and *hopZ5*. The gene encoding the T3E AvrRpt2 is present in all VCZ strains that have a narrow host range on cucurbits (Djitro et al. 2022a; Lacault et al. 2020) and in strains of the pathovar *tomato* (Innes et al. 1993), and was found in some strains of pathovars *maculicola*, *spinaceae*, and *tagetis* that were not known previously to have this gene. Indeed, *avrRpt2* was not included in the large analysis of the distribution of virulence encoding genes, including 53 T3Es, in *Pseudomonas syringae* strains made by Sarkar et al. (2006). BLASTn confirmed that those strains have the complete gene sequence in their genomes. As none of these strains, including those of the pathovar *tomato*, gave a positive signal in silico with the qPCR test for *sylC*, and neither experimentally for the strain of pathovar *tomato*, their reaction pattern using our multiplex test will be different to that of VCZ strains. The gene encoding the T3E *hopZ5* is present in VCZ strains that have a wide cucurbit host range and in the pathovar *actinidiae*, in which this T3E was described previously (Jayaraman et al. 2017). Strains of the pathovar *actinidiae* were not detected with the *sylC* qPCR assay, and hence gave a different result than VCZ strains with our multiplex test. We tried to improve the information provided by the test with the addition of qPCR tests for strains in clades 2b and 2d. However, only the test for clade 2d showed an efficiency consistent with that of the multiplex assay. We included this clade-2d qPCR test in the multiplex assay to identify cucurbit-associated strains, gain insight into their occurrence, and verify if they represent a low risk for zucchini seed production and distribution.

Specificity of the multiplex qPCR test was based on the combined results of different tests that were not individually specific, except the qPCR test designed to detect clade-2d strains. To be perfectly reliable, this multiplex qPCR test should be performed using isolated cultures of bacteria because, with crude seed extracts, false positive signals could arise from a mixture of strains. Such a warning was stated previously for a multiplex PCR test based on the simultaneous detection of two genes encoding T3Es dedicated to the identification of bacteria that cause common bacterial blight of bean (Boureau et al. 2013). However, the quantification of each marker that is possible with the qPCR assay could help solve this problem and allow the determination of whether marker-corresponding populations are compatible with single or mixed isolates. Furthermore, subsampling, which is performed to determine seed lot infection rates (ISTA 2015), could be used to help assess if the sample contains a single bacterial strain or mixed infections according to the occurrence of the strains simultaneously in the same subsamples or separately in different subsamples, as illustrated in Table 7.

More generally, detection tests based on T3Es or other virulence factors present de facto a risk of nonspecificity because these pathogenicity factors are easily transferable among strains, the repertoire of which evolves rapidly (Baltrus et al. 2011; Dillon et al. 2019a). Recombination and horizontal gene transfers were highlighted by several comparative genomic studies within the *P. syringae* species complex (Baltrus et al. 2017; Dillon et al. 2019b; Hulin et al. 2018; Newberry et al. 2019). However, targeting genes involved in pathogenicity could give clues to discriminate strains according to their interactions with plants, as was shown for *P. syringae* (Sarkar et al. 2006) and *Xanthomonas* (Hajri et al. 2009), but not for *Ralstonia solanacearum* (Peeters et al. 2013). For example, the PCR test based on the gene encoding SyrD, which is involved in the secretion of syringomycin and syringopeptin, is used to discriminate strains of the pathovar *syringae* from those of pathovar *morsprunorum* on cherry trees or from those of pathovar *pisi* on pea (Bultreys and Gheysen 1999). Another example is the harpin HrpZ, whose encoding gene is the preferred target for the detection of the pathovar *tomato*. Primers are designed in specific regions within this gene that is distributed among many pathovars. This gene is still in use to develop new molecular diagnosis tools (Chai et al. 2020; Chen et al. 2020; Zaccardelli et al. 2005).

While PCR is a highly sensitive and specific molecular biology technology widely used in the detection of various bacterial pathogens, it also detects intact DNA from dead bacteria. This is a serious problem when dealing with plant material such as seeds that could have been treated by different means and products (e.g., thermotherapy, disinfectants, or fungicides) that can affect bacterial viability but leave DNA detectable. Various alternative methods have been proposed to amplify DNA specifically from living cells. For example, the bioPCR assay, which includes a pre-enrichment step by cultivation on agar media, enables detection of viable and culturable bacteria (Schaad et al. 1995). However, using such an approach prevents quantification of initial bacterial population sizes. The use of DNA intercalating molecules such as ethidium monoazide (EMA) or propidium monoazide (PMA) enables differentiation of DNA between live and dead or membrane-damaged bacteria and enables detection of DNA only from viable cells (Chai et al. 2020). This latter approach is particularly interesting as it enables detection of DNA from bacterial cells that have been induced into a viable but nonculturable state (VBNC) by a stress, such as $CuSO_4$ treatment and low pH conditions in the case of the seedborne tomato pathogen *Clavibacter michiganensis* (Jiang et al. 2016). These intercalating dyes can be used in TaqMan qPCR assays to monitor bacterial pathogens selectively (e.g., Temple et al. 2013; Tian et al. 2016).

The protocol we used to test seeds for VCZ strains was adapted to discard PCR inhibitors present in zucchini seed extracts. The LOD of the multiplex qPCR test for spiked seed samples was as low as 2.5×10^3 CFU ml^{-1}, corresponding to 7.5×10^3 CFU g^{-1} of seeds. The DNA extraction protocol included a first step of centrifugation, followed by several purification steps and elution in a smaller volume than the initial volume to concentrate the samples (33-fold enrichment) without concentrating PCR inhibitors. The latter is important for PCR assays because sensitivity of the assays can be limited by the amount of sample that can be analyzed. Efficiencies of the qPCR tests developed in this study were as effective with spiked seed extracts after DNA extraction as the bacterial suspensions. The adaptation of the DNA extraction protocol is crucial to the success of the tests because the presence of polymerase inhibitors from plant material can have negative repercussions on assay performance, and not all protocols are adapted to work with different types of matrices or substrates (Lau and Botella 2017). For example, for the same DNA extraction protocol, the efficiency of detection of *Xylella fastidiosa* in planta varied depending on the plant matrix, probably due to abundant inhibitors such as polyphenols and polysaccharides. The negative impact may decrease by 10 to 100 times the LOD of the test when used with plant material in comparison with bacterial suspensions (Dupas et al. 2019).

The multiplex qPCR test developed in this study could be useful for epidemiological purposes to improve our knowledge of VCZ and constitutes an informative complement to the seed sector for the qPCR test currently in use that is based on *sylC*. It was possible in one step to identify different VCZ strains, to quantify them, and, by means of subsampling, to determine the infection rates of the seed lots by these strains. Compared to pre-existing methods (qPCR assay based on *sylC*, strain isolation in positive samples, and MLSA characterization), this multiplex qPCR

assay is particularly fast and cost effective. It could be useful for epidemiological studies aiming at identifying the sources of inoculum in parental seed lots or in the environment of seed production fields. The multiplex qPCR test will be very useful for seed testing because it provides information on the risk associated with infected seed lots. Depending on whether the seeds are infected with a strain with *avrRpt2* or *hopZ5*, the risk of losses to surrounding crops could be limited to *Cucurbita* spp. or extended to other cucurbits, such as melon, cucumber, and watermelon, especially since epidemics involving various cucurbits crops have been reported in the United States for closely related strains (Newberry et al. 2016, 2019). Precise identification of the strains responsible for epidemics is particularly relevant for strains with a wide host range, such as clade-2b strains belonging to the highly polyphagous pathovar *syringae* (Young 2010), or to the pathovar *aptata*, infecting both Cucurbitaceae and Chenopodiaceae (Koike et al. 2003; Morris et al. 2000; Nikolić et al. 2023). This would help identify sources of inoculum and trace possible infection in fields. The subsampling of seed lots enables detection and quantification of mixed infections. The detection limit of this multiplex qPCR assay is 5×10^5 CFU per sample of 500 seeds (for a 1,000-seed weight of 145 g). It would be interesting to determine the epidemiological risk associated with this detection threshold, for example, by a survey of fields sown with infected seed lots, and would inform the need to improve sensitivity of this detection tool for seed testing.

Acknowledgments

We thank P. Portier and C. Dutrieux (CIRM-CFBP) for strain preservation and J. Gouzy (CATI BARIC) for the genome assembly pipeline. We are also grateful to the SFR Quasav and the ANAN facility, in Angers, France, for DNA extraction and qPCR experiments.

Literature Cited

Amrein, H., Makart, S., Granado, J., Shakya, R., Schneider-Pokorny, J., and Dudler, R. 2004. Functional analysis of genes involved in the synthesis of syringolin A by *Pseudomonas syringae* pv. *syringae* B301D-R. Mol. Plant-Microbe Interact. 17:90-97.

Baltrus, D. A., McCann, H. C., and Guttman, D. S. 2017. Evolution, genomics and epidemiology of *Pseudomonas syringae*: Challenges in bacterial molecular plant pathology. Mol. Plant Pathol. 18:152-168.

Baltrus, D. A., Nishimura, M. T., Romanchuk, A., Chang, J. H., Mukhtar, M. S., Cherkis, K., Roach, J., Grant, S. R., Jones, C. D., and Dangl, J. L. 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog. 7:e1002132.

Bereswill, S., Bugert, P., Volksch, B., Ullrich, M., Bender, C. L., and Geider, K. 1994. Identification and relatedness of coronatine-producing *Pseudomonas syringae* pathovars by PCR analysis and sequence determination of the amplification products. Appl. Environ. Microbiol. 60: 2924-2930.

Berge, O., Monteil, C. L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D. C., and Morris, C. E. 2014. A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. PLoS One 9:e105547.

Borschinger, B., Bartoli, C., Chandeysson, C., Guilbaud, C., Parisi, L., Bourgeay, J. F., Buisson, E., and Morris, C. E. 2016. A set of PCRs for rapid identification and characterization of *Pseudomonas syringae* phylogroups. J. Appl. Microbiol. 120:714-723.

Boureau, T., Kerkoud, M., Chhel, F., Hunault, G., Darrasse, A., Brin, C., Durand, K., Hajri, A., Poussier, S., Manceau, C., Lardeux, F., Saubion, F., and Jacques, M.-A. 2013. A multiplex-PCR assay for identification of the quarantine plant pathogen *Xanthomonas axonopodis* pv. *phaseoli*. J. Microbiol. Methods 92:42-50.

Briand, M., Bouzid, M., Hunault, G., Legeay, M., Fischer-Le Saux, M., and Barret, M. 2021. A rapid and simple method for assessing and representing genome sequence relatedness. Peer Community J. 1:e24.

Briand, M., Gaborieau, R., Jacques, M.-A., Barret, M., Boureau, T., Gaillard, S., and Chen, N. W. G. 2016. SkIf: A tool for rapid identification of genes or regulators of interest. F1000research 5:1593.

Bull, C. T., and Koike, S. T. 2015. Practical benefits of knowing the enemy: Modern molecular tools for diagnosing the etiology of bacterial diseases and understanding the taxonomy and diversity of plant-pathogenic bacteria. Annu. Rev. Phytopathol. 53:157-180.

Bultreys, A., and Gheysen, I. 1999. Biological and molecular detection of toxic lipodepsipeptide-producing *Pseudomonas syringae* strains and PCR identification in plants. Appl. Environ. Microbiol. 65:1904-1909.

Chai, A.-l., Ben, H.-y., Guo, W.-t., Shi, Y.-x., Xie, X.-w., Li, L., and Li, B.-j. 2020. Quantification of viable cells of *Pseudomonas syringae* pv. *tomato* in tomato seed using propidium monoazide and a real-time PCR assay. Plant Dis. 104:2225-2232.

Chen, Z.-D., Kang, H.-J., Chai, A.-L., Shi, Y.-X., Xie, X.-W., Li, L., and Li, B.-J. 2020. Development of a loop-mediated isothermal amplification (LAMP) assay for rapid detection of *Pseudomonas syringae* pv. *tomato* in planta. Eur. J. Plant Pathol. 156:739-750.

Denancé, N., Briand, M., Gaborieau, R., Gaillard, S., and Jacques, M.-A. 2019. Identification of genetic relationships and subspecies signatures in *Xylella fastidiosa*. BMC Genom. 20:239.

Dillon, M. M., Almeida, R. N. D., Laflamme, B., Martel, A., Weir, B. S., Desveaux, D., and Guttman, D. S. 2019a. Molecular evolution of *Pseudomonas syringae* type III secreted effector proteins. Front. Plant Sci. 10:418.

Dillon, M. M., Thakur, S., Almeida, R. N. D., Wang, P. W., Weir, B. S., and Guttman, D. S. 2019b. Recombination of ecologically and evolutionarily significant loci maintains genetic cohesion in the *Pseudomonas syringae* species complex. Genome Biol. 20:3.

Djitro, N., Roach, R., Mann, R., Campbell, P. R., Rodoni, B., and Gambley, C. 2022a. Comparative genomics and genomic diversity of *Pseudomonas syringae* clade 2b-a in Australia. BMC Microbiol. 22:278.

Djitro, N., Roach, R., Mann, R., Rodoni, B., and Gambley, C. 2022b. Characterization of *Pseudomonas syringae* isolated from systemic infection of zucchini in Australia. Plant Dis. 106:541-548.

Dudnik, A., and Dudler, R. 2014. Genomics-based exploration of virulence determinants and host-specific adaptations of *Pseudomonas syringae* strains isolated from grasses. Pathogens 3:121-148.

Dupas, E., Briand, M., Jacques, M.-A., and Cesbron, S. 2019. Novel tetraplex quantitative PCR assays for simultaneous detection and identification of *Xylella fastidiosa* subspecies in plant tissues. Front. Plant Sci. 10:1732.

European and Mediterranean Plant Protection Organization. 2019. PM 7/98 (4) Specific requirements for laboratories preparing accreditation for a plant pest diagnostic activity. EPPO Bull. 49:530-563.

Gallelli, A., Talocci, S., Pilotti, M., and Loreti, S. 2014. Real-time and qualitative PCR for detecting *Pseudomonas syringae* pv. *actinidiae* isolates causing recent outbreaks of kiwifruit bacterial canker. Plant Pathol. 63: 264-276.

Gomila, M., Busquets, A., Mulet, M., García-Valdés, E., and Lalucat, J. 2017. Clarification of taxonomic status within the *Pseudomonas syringae* species group based on a phylogenomic analysis. Front. Microbiol. 8:2422.

Guilbaud, C., Morris, C. E., Barakat, M., Ortet, P., and Berge, O. 2016. Isolation and identification of *Pseudomonas syringae* facilitated by a PCR targeting the whole *P. syringae* group. FEMS Microbiol. Ecol. 92: fiv146.

Hajri, A., Brin, C., Hunault, G., Lardeux, F., Lemaire, C., Manceau, C., Boureau, T., and Poussier, S. 2009. A "repertoire for repertoire" hypothesis: Repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. PLoS One 4:e6632.

Hulin, M. T., Armitage, A. D., Vicente, J. G., Holub, E. B., Baxter, L., Bates, H. J., Mansfield, J. W., Jackson, R. W., and Harrison, R. J. 2018. Comparative genomics of *Pseudomonas syringae* reveals convergent gene gain and loss associated with specialization onto cherry (*Prunus avium*). New Phytol. 219:672-696.

Hwang, M. S. H., 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.

Innes, R. W., Bent, A. F., Kunkel, B. N., Bisgrove, S. R., and Staskawicz, B. J. 1993. Molecular analysis of avirulence gene *avrRpt2* and identification of a putative regulatory sequence common to all known *Pseudomonas syringae* avirulence genes. J. Bacteriol. 175:4859-4869.

International Seed Testing Association (ISTA). 2015. Procedure organizing and analyzing results of the seed health proficiency tests. https://www.seedtest.org/api/rm/ME8E23XMB8XA692/pt-p-3organizingandanalyzingshpts-v2-0-2.pdf

Jacques, M.-A., Durand, K., Orgeur, G., Balidas, S., Fricot, C., Bonneau, S., Quillévéré, A., Audusseau, C., Olivier, V., Grimault, V., and Mathis, R. 2012. Phylogenetic analysis and polyphasic characterization of *Clavibacter michiganensis* strains isolated from tomato seeds reveal that nonpathogenic strains are distinct from *C. michiganensis* subsp. *michiganensis*. Appl. Environ. Microbiol. 78:8388-8402.

Jayaraman, J., Choi, S., Prokchorchik, M., Choi, D. S., Spiandore, A., Rikkerink, E. H., Templeton, M. D., Segonzac, C., and Sohn, K. H. 2017. A bacterial acetyltransferase triggers immunity in *Arabidopsis thaliana* independent of hypersensitive response. Sci. Rep. 7:3557.

Jiang, N., Lv, Q. Y., Xu, X., Cao, Y. S., Walcott, R. R., Li, J. Q., and Luo, L. X. 2016. Induction of the viable but nonculturable state in *Clavibacter michiganensis* subsp. *michiganensis* and *in planta* resuscitation of the cells on tomato seedlings. Plant Pathol. 65:826-836.

Khalaf, E. M., and Raizada, M. N. 2016. Taxonomic and functional diversity of cultured seed associated microbes of the cucurbit family. BMC Microbiol. 16:131.

Koike, S. T., Henderson, D. M., Bull, C. T., Goldman, P. H., and Lewellen, R. T. 2003. First report of bacterial leaf spot of Swiss chard caused by *Pseudomonas syringae* pv. *aptata* in California. Plant Dis. 87:1397.

Koressaar, T., and Remm, M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289-1291.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol. Biol. Evol. 33:1870-1874.

Lacault, C., Briand, M., Jacques, M.-A., and Darrasse, A. 2020. Zucchini vein clearing disease is caused by several lineages within *Pseudomonas syringae* species complex. Phytopathology 110:744-757.

Lamichhane, J. R., Fabi, A., and Varvaro, L. 2014. Bacterial species associated to brown spots of hazelnut in Central Italy: Survey, isolation and characterization. Acta Hortic. 1052:243-251.

Lamichhane, J. R., Messéan, A., and Morris, C. E. 2015. Insights into epidemiology and control of diseases of annual plants caused by the *Pseudomonas syringae* species complex. J. Gen. Plant Pathol. 81:331-350.

Langston, D. B., Jr., Sanders, F. H., Brock, J. H., Gitaitis, R. D., Flanders, J. T., and Beard, G. H. 2003. First report of a field outbreak of a bacterial leaf spot of cantaloupe and squash caused by *Pseudomonas syringae* pv. *syringae* in Georgia. Plant Dis. 87:600.

Lau, H. Y., and Botella, J. R. 2017. Advanced DNA-based point-of-care diagnostic methods for plant diseases detection. Front. Plant Sci. 8: 2016.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D. W., Yiu, S.-M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T.-W., and Wang, J. 2012. SOAPdenovo2: An empirically improved memory-efficient short-read *de novo* assembler. GigaScience 1:18.

Manceau, C., Gironde, S., Briand, B., and Lybeert, H. 2011. Means and methods for detecting and identifying a novel bacterium responsible for phytosanitary disorders in plants (zucchini) and novel resistant plants. Patent #WO2011003984, World Intellectual Property Organization, Geneva, Switzerland.

Maury, Y., Duby, C., Bossennec, J.-M., and Boudazin, G. 1986. Group analysis using ELISA: Determination of the level of transmission of soybean mosaic virus in soybean seed. Agronomie 5:405-415.

Meng, X.-L., Xie, X.-W., Shi, Y.-X., Chai, A.-L., Ma, Z.-H., and Li, B.-J. 2017. Evaluation of a loop-mediated isothermal amplification assay based on *hrpZ* gene for rapid detection and identification of *Pseudomonas syringae* pv. *lachrymans* in cucumber leaves. J. Appl. Microbiol. 122: 441-449.

Morris, C. E., Glaux, C., Latour, X., Gardan, L., Samson, R., and Pitrat, M. 2000. The relationship of host range, physiology, and genotype to virulence on cantaloupe in *Pseudomonas syringae* from cantaloupe blight epidemics in France. Phytopathology 90:636-646.

Morris, C. E., Monteil, C. L., and Berge, O. 2013. The life history of *Pseudomonas syringae*: Linking agriculture to earth system processes. Annu. Rev. Phytopathol. 51:85-104.

Newberry, E. A., Ebrahim, M., Timilsina, S., Zlatković, N., Obradović, A., Bull, C. T., Goss, E. M., Huguet-Tapia, J. C., Paret, M. L., Jones, J. B., and Potnis, N. 2019. Inference of convergent gene acquisition among *Pseudomonas syringae* strains isolated from watermelon, cantaloupe, and squash. Front. Microbiol. 10:270.

Newberry, E. A., Jardini, T. M., Rubio, I., Roberts, P. D., Babu, B., Koike, S. T., Bouzar, H., Goss, E. M., Jones, J. B., Bull, C. T., and Paret, M. L. 2016. Angular leaf spot of cucurbits is associated with genetically diverse *Pseudomonas syringae* strains. Plant Dis. 100:1397-1404.

Nikolić, I., Glatter, T., Ranković, T., Berić, T., Stanković, S., and Diepold, A. 2023. Repertoire and abundance of secreted virulence factors shape the pathogenic capacity of *Pseudomonas syringae* pv. *aptata*. Front. Microbiol. 14:1205257.

Peeters, N., Carrère, S., Anisimova, M., Plener, L., Cazalé, A.-C., and Genin, S. 2013. Repertoire, unified nomenclature and evolution of the Type III effector gene set in the *Ralstonia solanacearum* species complex. BMC Genom. 14:859.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. 2020. Using SPAdes *de novo* assembler. Curr. Protoc. Bioinformatics 70:e102.

Prosen, D., Hatziloukas, E., Schaad, N. W., and Panopoulos, N. J. 1993. Specific detection of *Pseudomonas syringae* pv. *phaseolicola* DNA in bean seed by polymerase chain reaction-based amplification of a phaseolotoxin gene region. Phytopathology 83:965-970.

Qiao, P., Zhao, M., Liu, D., Guan, W., Bai, Q., Yang, Y., and Zhao, T. 2023. Identification of the causal agent of bacterial leaf spot on watermelon in China. Plant Dis. 107:1839-1846.

Sarkar, S. F., Gordon, J. S., Martin, G. B., and Guttman, D. S. 2006. Comparative genomics of host-specific virulence in *Pseudomonas syringae*. Genetics 174:1041-1056.

Schaad, N. W., Cheong, S. S., Tamaki, S., Hatziloukas, E., and Panopoulos, N. J. 1995. A combined biological and enzymatic amplification (BIO-PCR) technique to detect *Pseudomonas syringae* pv. *phaseolicola* in bean seed extracts. Phytopathology 85:243-248.

Sedighian, N., Shams-Bakhsh, M., Osdaghi, E., and Khodaygan, P. 2014. Etiology and host range of bacterial leaf blight and necrosis of squash and muskmelon in Iran. J. Plant Pathol. 96:507-514.

Swaroop, S. 1951. The range of variation of the most probable number of organisms estimated by the dilution method. Ind. J. Med. Res. 39: 107-134.

Temple, T. N., du Toit, L. J., Derie, M. L., and Johnson, K. B. 2013. Quantitative molecular detection of *Xanthomonas hortorum* pv. *carotae* in carrot seed before and after hot-water treatment. Plant Dis. 97:1585-1592.

Tian, Q., Feng, J.-j., Hu, J., and Zhao, W.-j. 2016. Selective detection of viable seed-borne *Acidovorax citrulli* by real-time PCR with propidium monoazide. Sci. Rep. 6:35457.

Xin, X.-F., Kvitko, B., and He, S. Y. 2018. *Pseudomonas syringae*: What it takes to be a pathogen. Nat. Rev. Microbiol. 16:316-328.

Young, J. M. 2010. Taxonomy of *Pseudomonas syringae*. J. Plant Pathol. 92:S1.5-S1.14.

Zaccardelli, M., Spasiano, A., Bazzi, C., and Merighi, M. 2005. Identification and *in planta* detection of *Pseudomonas syringae* pv. *tomato* using PCR amplification of $hrpZ_{Pst}$. Eur. J. Plant Pathol. 111:85-90.

Zerbino, D. R., and Birney, E. 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res. 18:821-829.