Zucchini vein clearing disease is caused by several lineages within Pseudomonas syringae species complex.

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

Zucchini vein clearing disease is caused by several lineages within *Pseudomonas syringae* species complex.

Caroline Lacault [1], Martial Briand [1], Marie-Agnès Jacques [1] and Armelle Darrasse [1]

[1] IRHS, Agrocampus-Ouest, INRA, Université d'Angers, SFR 4207 QuaSaV, 49071, Beaucouzé, France

Corresponding authors: Armelle Darrasse, E-mail address: armelle.darrasse@inra.fr; Marie-Agnès Jacques, E-mail address : marie-agnes.jacques@inra.fr

Abstract

Zucchini (*Cucurbita pepo*) is worldwide affected by *Pseudomonas syringae* inducing vein clearing, stunting and necroses during plantlet development. A collection of 58 *P. syringae* strains isolated from diseased zucchini plantlets was characterized by multilocus sequence analysis (MLSA). A subset of 23 strains responsible for vein clearing of zucchini (VCZ) was evaluated for pathogenicity on zucchini and their genomes were sequenced. Host range of six VCZ strains was evaluated on 11 cucurbit species. Most VCZ strains belong to clades 2a and 2b-a within phylogroup 2 of *P. syringae* species complex and are closely related to other strains previously isolated from cucurbits. Genome analyses revealed diversity among VCZ strains within each clade. One main cluster, once referred to by the invalid pathovar name [peponis], gathers VCZ strains presenting a narrow host range including zucchini and squashes. Other VCZ strains present a large host range including zucchini, squashes, cucumber, melons and in some cases watermelon. The VCZ strain pathogenic features are strongly associated with type III effector repertoires. Presence of *avrRpt2* and absence of *hopZ5* are associated with a narrow host range, while presence of *hopZ5* and absence of *avrRpt2* are most generally associated with a large host range. In order to better detect the different clusters identified with whole genome sequence and pathogenicity analyses, we used a specific-k-mers approach to refine the MLSA scheme. Using this novel MLSA scheme to type *P. syringae* isolates from diseased cucurbits would give insight into distribution of worldwide strains and origin of epidemics.

Additional key-words: seedborne disease, bacteria

Key-words: *Pseudomonas syringae*, *Cucurbita pepo*, cucurbits, bacterial seedborne disease

Introduction

In 2004 diseased plantlets of zucchini (*Cucurbita pepo* subsp. *pepo*) were observed in plant grower's facilities in Europe. Diseased plants show cotyledon and leaf necroses, vein clearing and stunting. This disease is observed on plantlets and impacts plant development at early stages but to date no symptoms are observed on fruits. One causal agent has been identified as belonging to the *Pseudomonas syringae* species complex and was provisionally named *P. syringae* pv. *peponis* (Manceau et al. 2011). This pathogenic bacterium was shown to be seed-transmitted (Manceau et al. 2011). Climatic conditions drastically influence symptom expression at early stage development, for example low temperatures are associated with increased disease severity. Because most generally efficient chemical options are not available toward bacterial threats, gain of knowledge about *P. syringae* responsible for vein clearing on zucchini (VCZ) is essential to implement appropriate control strategies.

The *P. syringae* species complex comprises 19 phylogenomic species based on Average Nucleotide Identity (ANI) (Gomila et al. 2017), some of which are formally described (*P. amygdali*, *P. asturiensis*, *P. avellanae*, *P. cannabina*, *P. caricapapayae*, *P. caspiana*, *P. cerasi, P. cichorii*, *P. congelans*, *P. syringae*, and *P. viridiflava*). Moreover, *P. coronafaciens* has been proposed recently by Dutta and colleagues (2018). However, some of the crop pathogenic species are not easy to refer to as recombination between lineages is sufficiently high to blur the taxonomic boundaries (Baltrus et al. 2017; Dillon et al. 2019a). Many studies devoted to classify pathogenic strains of *P. syringae sensu lato* refer to phylogroups. Thirteen phylogroups (i.e some phylogroups cluster several phylogenomic species) generally subdivided into clades were defined in the *P. syringae*

species complex based on Multi Locus Sequence Analysis (MLSA) (Berge et al. 2014) and over 60 pathovars were described based on pathogenicity (Young 2010). Indeed, a pathovar is defined as a group of strains showing the same or similar characteristics at an infra-subspecific level but presenting a distinctive pathogenicity, which means either a unique host range and/or unique symptomatology (Dye et al. 1980).

Various pathogenic *P. syringae sensu lato* have been reported on cucurbits. *P. amygdali* pv. *lachrymans* causes angular leaf spots on cucumber (Bradburry 1986) and other cucurbits (Hopkins and Schenck 1972; Shila et al. 2013). Other strains of *P. syringae* that were named pathovar *lachrymans* cause different symptoms on cucumber and belong to the phylogenomic species *P. tomato* within the phylogroup 1 (Gomila et al. 2017; Słomnicka et al. 2018). Diverse strains isolated from melons and squashes within phylogroup 2 belong to *P. syringae* pv. *aptata* that was initially described as pathogenic on sugar beet (Morris et al. 2000; Sedighian et al. 2014). Other strains of *P. syringae sensu lato* genetically different from the two previously mentioned pathovars were shown to induce bacterial leaf spots on adult watermelon, melon and squash (Newberry et al. 2016, 2017, 2018, and 2019; Tian et al. 2017) and bacterial warts on pumpkin fruits (Tymon and Inglis 2017). Most of these cucurbit pathogenic strains are phylogenetically diverse (Newberry et al. 2016, 2017, 2018, and 2019). These strains belong to different species including *P. cerasi* (Kaluzna et al. 2016) that corresponds to phylogroup 2a, (Berge et al. 2014), *P. syringae* (corresponding to phylogroup 2b, Berge et al. 2014), the unnamed group A *sensu* Gomila *et al.* (2017; also named phylogroup 2d, Berge et al. 2014) and the phylogroup 2b-a issued from the recombination of *P. syringae* and *P. cerasi* (Newberry et al. 2019). In other words, these strains belong to different clades, i.e. 2a, 2b,

2b-a and 2d, within the phylogroup 2 of the *P. syringae* species complex. The phylogroup 2 appears to be the most ubiquitous group and comprises numbers of pathogenic strains with narrow or wide host range but also strains recovered from the environment pathogenic or not (Baltrus et al. 2017; Berge et al. 2014; Monteil et al. 2013; Morris et al. 2000).

Cucurbitaceae is a family of plants including economically important crops such as cucumbers (*Cucumis sativus* L.), melons (*Cucumis melo* L.), watermelons (*Citrullus lanatus*) and squashes. Squash is a generic name for various plant species including *C. pepo*, *C. maxima* and *C. moschata*. Zucchini together with pumpkin, vegetable marrow and cocozelle form *C. pepo* subsp. *pepo.* Acorn, scallop, straightneck and crookneck form *C. pepo* subsp. *texana* (Paris 1989; Paris et al. 2015). Genetic relationships between many *C. pepo* were inferred from the calculation of average dissimilarity values using DNA markers (Paris et al. 2015). These results indicate that zucchini and pumpkin from the same subspecies are genetically close while zucchini and straightneck are more distant.

Pathogenicity assays are essential to determine epidemic potential of phytopathogenic bacteria. Such assays constitute a prerequisite to designate pathovars distinguished by distinct host range and symptoms (Bull et al. 2008). Host range and virulence of *P. syringae sensu lato* are shaped by virulence factors such as Type III effectors (T3Es) and phytotoxins (Silby et al. 2011; Xin et al. 2018). T3Es are secreted via the Type III secretion system (T3SS) to specifically target host proteins or DNA sequences inside plant cells and alter physiological processes to the benefit of the pathogens (Feng and Zhou 2012). The *P. syringae* content of T3Es evolves and T3Es can be acquired via horizontal transfer as response to host-mediated selection (Baltrus et al. 2011; Dillon et al. 2019a, 2019b).

Several methods based on genomic DNA sequences are used to study organism phylogeny and update taxonomy. MLSA is very useful to assign strains to previously described taxa and reveal phylogenetic relationships among strains (Jacques et al. 2012, Jacques et al. 2016 , Young et al. 2010). This portable and easy to use method is based on four or more housekeeping genes. Within the *P. syringae* species complex, a classical MLSA scheme initially proposed by Sarkar and Guttman (2004) and refined by Hwang et al. (2005) is frequently used to describe genetic structure of strain groups and to assign unknown strains to the correct genetic group and pathovar within the *P. syringae* species complex (Bull et al. 2011; Cunty et al. 2015; Ferrante and Scortichini 2015; Martin-Sanz et al. 2013; Newberry et al. 2016; Słomnicka et al. 2015). Furthermore, several methods based on whole genome sequence comparisons are currently used to infer closely related bacteria relationships, such as the average nucleotide identity based on BLAST (ANIb) (Konstantinidis and Tiedje 2005). Other methods, based on the number of words of length k (*k*-mers) shared between genome sequences, could be used as proxy for species delimitation (Briand et al. 2019). In addition, distribution of *k*-mers in specific genomic group could provide a list of markers for the identification of bacterial strains (Denancé et al. 2019).

In order to gain knowledge on bacteria causing VCZ and to define appropriate field management strategies, we developed a three-step approach. Our first aim was to correctly allocate the VCZ strains in the *P. syringae* species complex and infer how these strains are related to other *P. syringae* isolated from cucurbits with a MLSA and a whole-genome studies. Second, different pathogenicity assays were designed to finely phenotype strains on their host of origin (zucchini) but also on diverse cucurbit genotypes.

Third, based on draft genome sequences, some specific T3Es were highlighted for their potential link with host range specialization and a new MLSA scheme dedicated to this group of pathogens was developed. This work allows for better evaluation of the epidemic potential of VCZ strains.

Materials and methods

Bacterial strains collection. Fifty-eight strains were isolated from zucchini seedlings (*C. pepo* subsp. *pepo*) with symptoms of vein clearing issued from seed lots harvested in different geographic locations between 2002 and 2015 (Table 1). Bacterial strains were stored in a -80°C freezer in 40% glycerol solution and were routinely cultivated at 28°C on $TSA_{10\%}$ media (3 g/L Tryptone Soya broth, 15 g/L Agar). All these strains induced typical VCZ symptoms (necrosis, vein clearing and stunting) on zucchini following inoculation by rubbing the cotyledons (see the description below in pathogenicity test section).

Housekeeping gene sequencing. Amplicons were obtained from fresh pure cultures (overnight growth at 28°C on $TSA_{10\%}$) using primers developed by Hwang and collaborators (2005) for the four following housekeeping genes, *gapA* (encoding the glyceraldehyde-3-phosphate dehydrogenase A), *gltA* (encoding the citrate synthase), *gyrB* (encoding the B subunit of DNA gyrase) and *rpoD* (encoding the RNA polymerase sigma70 factor RpoD). Briefly, PCRs were performed in a 50 µL reaction mixture containing 0.2 µL 5X GoTaq reaction buffer (Promega), 0.2 µM dNTP, 0.5 µM of each primer and 10 µL of boiled bacterial cells. The amplification program consisted of 35 cycles of denaturation at 94°C for 30 s, annealing at 55 °C (*gapA*) or 64°C (*gltA, gyrB, rpoD*) for 30 s, extension at 72 °C for 1 min 30 s and a final extension for 7 min at 72°C. Sanger sequencing (Genoscreen, Lille, France) was performed on forward and reverse strands.

Sequence acquisition and alignment. Forward and reverse nucleotide sequences were assembled and aligned according to the reading frame of a complete gene sequence and trimmed with Bionumerics, version 7.6 (Applied Maths, Sint-Martens-Latem, Belgium) and Geneious, version 9.1.7 (https://www.geneious.com). Sequences were trimmed to obtain fragments of the following size: 474 bp (*gapA*), 528 bp (*gltA*), 507 bp (*gyrB*) and 498 pb (*rpoD*) and eventually concatenated according to the alphabetical order in a 2007 bp long sequence.

MLSA. Phylogenetic analyses were first performed on individual partial gene sequences of *gapA, gltA, gyrB* and *rpoD* and then on the concatenated dataset to obtain one single fragment of 2007 bp. These sequences were analyzed together with publicly available sequences from 22 strains representing the main phylogroups and clades (Berge et al. 2014) of the *P. syringae* species complex and from 17 *P. syringae* strains (13-509A, Ps711, 13-140A, 13-C2, 14-32, 14-410, 14-Gil, Zum3584, 200-1, Bs2121, Zum3984, 03-19A, Bs13-139B, 13-429, Ps03-A13, PsBS3058 and Ps02-B3) isolated from various cucurbits (Newberry et al. 2016, 2019) (Supplementary table 1). Sequences from strain CFBP 2107 (*Pseudomonas viridiflava*) were used to root the trees.

For the novel MLSA scheme proposed in this study, the analysis of the four gene sequences previously mentioned (*gapA, gltA, gyrB, rpoD*) and genes identified in this study *Psyr3420, Psyr4880* and *Psyr3208* (see results section) lead to the acquisition of one unique fragment of 4431 bp. Sequences were recovered from the genome sequences of a strain subset from our collection (23 VCZ strains), a subset of 14 strains isolated from cucurbits by Newberry et al. (2019) and 22 diverse strains from phylogroup 2 (Supplementary table 1).

The Shimodaira-Hasegawa (SH) (Shimodaira and Hasegawa 1999) test implemented in the DnaML program from PHYLIP, Phylogeny Inference Package, version 3.69 (Felsenstein 1989) was used to test whether tree topologies fell within the same confidence limits. Maximum-likelihood (ML) trees were generated with MEGA 7 (Kumar et al. 2016) using the Tamura-Nei model with gamma distributed with invariant sites (G+I), and 1,000 bootstrap replicates.

Genome sequencing. A subset of 23 VCZ strains were chosen to maximize their genetic diversity and their geographical origin (P99, P121, P123, P12855, P84, P12831, P139, P12857, P127, P79, P22, P12854, P12832, P113, P5, P118, P66, P77, P78, P87, P89, P73 and P108). The pathotype strain of *P. syringae* pv. *dysoxyli* (strain CFBP 2356) was also included in this study. Genomic DNAs of these strains were obtained using the Wizard® Genomic DNA Purification Kit (Promega) according to the manufacturer's instructions. The genomic DNA quality was checked with NanoDrop ND-1000 spectrophotometer (the NanoDrop Technologies, Wilmington, DE) and quantified using a Qubit fluorometer (Invitrogen). Sequencing libraries were prepared using TruSeq DNA PCR-Free kits (Illumina). DNAs of P66 and P99 strains were paired-end sequenced (2 x 300 bp) using the Illumina MiSeq v3 (ANAN plateform, SFR QuaSav, Angers, France). DNAs of the other VCZ strains and strain CFBP 2356 were paired-end sequenced (2 x 150 bp) using the Illumina HiSeq X-Ten (Beijing Genomics Institute, China).

Genome sequence analysis. 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) Genome sequences were annotated using EuGene-PP (Sallet, Gouzy, and Schiex 2014). A dataset of 116 genome sequences (Supplementary table 1)

composed of 23 genome sequences from strains responsible for VCZ, a collection of 14 genome sequences (13-509A, Ps711, 13-140A, 13-C2, 14-32, 14-410, 14-Gil, Zum3584, 200-1, Bs2121, Zum3984, 03-19A, Bs13-139B and 13-429) of *P. syringae* isolated from cucurbits (Newberry et al. 2019), the genome sequence of CFBP 2356 (*P. syringae* pv. *dysoxyli*) and 78 genome sequences available in NCBI (to represent *P. syringae* diversity from phylogroup 2) was used for phylogenetic analyses. To assess relationships between genome sequences, the average nucleotide identity using blast (ANIb; Konstantinidis and Tiedje 2005) and percentage of shared 15-mers (Briand et al. 2019) were employed. ANIb values were obtained with the online version of pyani (https://gihub.com/widdowquinn/pyana) while percentage of shared *k*-mers and distance matrices were calculated with Ki-S as were dendrograms (https://iris.angers.inra.fr/galaxypub-cfbp) (Briand et al. 2019). Specific *k*-mers in predicted CDSs of each genomic group were identified with SkIf (Denancé et al. 2019). SkIf was also used to identify sequences shared by all VCZ strains and absent in the predicted proteins of other closely-related pathovars. Homology groups were identified in predicted proteomes by an OrthoMCL analysis (80% identity on 80% length) using Family companion tool (http://family-companion.toulouse.inra.fr) (Cottret et al. 2018). Finally, the T3E repertoire of each strain was predicted through tBLASTN search (identity higher than 80% on at least 80% of CDS length) on 138 amino acid sequences of T3Es (Hops database at http://www.pseudomonas-syringae.org/). Putative frameshift mutations were checked by aligning each T3E sequence with the corresponding curated reference sequence with MultAlin (http://multalin.toulouse.inra.fr/multalin/) (Corpet 1988).

Pathogenicity tests on zucchini. To characterize the pathogenicity of strains on their host of origin (zucchini), a cotyledon rubbing test and two seed soaking tests were run on *C. pepo* subsp. *pepo* cvs. Tosca and Opale. For each test, 23 strains responsible for VCZ (P99, P121, P123, P12855, P84, P12831, P139, P12857, P127, P79, P22, P12854, P12832, P113, P5, P118, P66, P77, P78, P87, P89, P73 and P108) and four controls were inoculated. Water was used as negative control, CFBP 5427 (*P. syringae* pv. *aptata*) and CFBP 6463 (*P. amygdali* pv. *lachrymans*) were used as pathogenic control on zucchini and CFBP 2356 (*P. syringae* pv. *dysoxyli*), which is a woody plant pathogen was used as negative control.

In the cotyledon rubbing test, five plantlets at the cotyledon stage were inoculated per strain and cultivar by rubbing three times the cotyledon surface with a finger soaked in bacterial suspensions calibrated at 1×10^8 cfu/mL. Plantlets were produced in a greenhouse from healthy seeds during seven days in pots containing Klassman substrate and watered as required for optimal growth. Inoculated plantlets were incubated in a growth chamber (18°C night, 25°C day, 14h day-light, relative humidity near 100 %) and the health status of each plantlet was recorded eight-days post-inoculation (dpi). Plantlets were qualitatively rated and the number of plantlets with necrosis on leaves, vein clearing on leaves and stunting was counted.

For the two seed soaking tests, 10 seeds per strain and per zucchini cultivar (*C. pepo* subsp. *pepo* cvs. Tosca, Opale) were inoculated by a 30 min soaking in a fresh bacterial suspension calibrated at 1×10^6 cfu/mL or 1×10^7 cfu/mL. Then, seeds were dried under a laminar flow in individual Petri dishes containing filter paper during 2 hours. Each dried seed was sown in an individual small pot containing Klassman substrate and was watered.

Pots were incubated in a climatic chamber (18°C night, 25°C day, 14h day-light, relative humidity near 100 %). Health status of each plantlet was evaluated eight days after sowing for both inoculum doses and up to 14 days for test with a 1×10^7 cfu/mL inoculum concentration. Plantlets were considered healthy when no macroscopic symptom was visible and diseased when plantlets presented at least one of the following symptoms: necrosis on cotyledons, vein clearing of the first leaf or stunting.

Pathogenicity tests on different cucurbit species. To evaluate the host range of the VCZ strains, pathogenicity tests were run on a collection of 11 cucurbit genotypes. Six squashes including zucchini (*Cucurbita pepo* subsp. *pepo* cvs. Opale and Tosca), a straightneck squash (*C. pepo* subsp. *texana* cv. Cheetah), a pumpkin (*C. pepo* subsp. *pepo* cv. *Mischief),* a French pumpkin (*Cucurbita maxima* cv. Rouge Vif d'Etampes) and a butternut (*Cucurbita moschata* cv. Sibelle) were tested. Three melon varieties including a cantaloupe (*Cucumis. melo* var. *cantalupsis* cv. Perseus), a Galia melon (*C. melo* var. *galia* cv. Citrex) and a Canary melon (*C. melo* var. *inodorus* cv. Helios), a cucumber (*Cucumis sativus* cv. Dahab) and a watermelon (*Cucumis lanatus* cv. Livia) were tested. Seeds were sown individually in pots containing Klassman substrate and germinated in a greenhouse until the first true leaf stage. The day before inoculation, plantlets were moved for acclimation to inoculation cages in growth chamber (18°C night, 25°C day during 14h). Each inoculation cage contained five plantlets of each cucurbit genotype. Six strains responsible for VCZ (P99, P79, P66, P73, P77, P108) and three controls including water, CFBP 5427 (*P. syringae* pv. *aptata*) and CFBP 6463 (*P. amygdali* pv. *lachrymans*) were evaluated. The assay was independently done twice for each genotype (except for *C. pepo* subsp. *texana* cv. Cheetah). Bacterial suspensions were prepared from fresh

cultures and calibrated at 1 × 10^7 cfu/mL. Each strain was sprayed until runoff on each plantlet. Plantlets were incubated seven days (18°C night, 25°C day during 14 h and near 100 % relative humidity). Pathogenicity was qualitatively rated as (0) when a plantlet was asymptomatic and (1) when plantlets presented necrosis. First leaves were collected for photography on a light table.

Test of hypersensitive response (HR). Ability of strains P99, P66, P108 and P12857 to elicit a HR was tested on melons (*C. melo* var. *galia* cv. Citrex, *C. melo* var. *cantalupsis* cv. Perseus, *C. melo* var. *inodorus* cv. Helios), cucumber (*C. sativus* cv. Dahab) and watermelon (*C. lanatus* cv. Livia). Additionally, strains P79, P77 and P73 were tested on watermelon. These strains were chosen based on phylogenetic position and to represent different pathogenicity groups. Pathogenicity controls were CFBP 5427 (*P. syringae* pv. *aptata*) on melons, CFBP 6463 (*P. amygdali* pv. *lachrymans*) on cucumber and CFBP 4459 (*Acidovorax citrulli*) on watermelon. HR control was given by strain CFBP 2356 (*P. syringae* pv. *dysoxyli*) as this strain is non-pathogenic on cucurbits. Water was used as a negative control. At the time of inoculation, fresh inocula calibrated at 1 × 10^8 cfu/mL were infiltrated with a needle in the limb of the second leaf of 3-weeks old plantlets produced in Klassman substrate in a greenhouse. The HR characteristics (rapid death of cells in the infiltration area) were visually evaluated at 48 and 72 h after inoculation on each inoculated leaf. Dried brown necrosis limited to the infiltrated area corresponded to HR-like while spreading lesions that eventually collapsed with a water-soaked aspect corresponded to disease.

Nucleotide sequence accession numbers. The GenBank accession numbers for the partial sequences used in this study are MN176747 to MN176804 for *gapA*; MN176805

to MN176862 for *gltA*; MN176863 to MN176920 for *gyrB* and MN176921 to MN176978 for *rpoD*. Genomic sequences of VCZ strains sequenced in this study were deposited in NCBI, accession numbers are listed in table 1. The genomic sequence of *P. syringae* pv. *dysoxyli* CFBP 2356 sequenced in this study was deposited in NCBI, accession number is VLHW00000000.

Results

Phylogenetic analyses of VCZ strains. All strains responsible for VCZ belonged to phylogroup 2 of the *P. syringae* species complex (Fig. 1A). The Maximum Likelihood tree clearly showed that all VCZ strains grouped in the same branch as did type strains and pathotype strains of phylogroup 2 (Fig. 1A). Other strains of *P. syringae sensu lato* isolated from various cucurbits (Newberry et al. 2016, 2019) also grouped in the same cluster. VCZ strains were genetically different from strain NM002 of *P. amygdali* pv. *lachrymans*, which is a pathovar previously reported as pathogenic on various cucurbits (Bradbury 1986) and belonging to phylogroups 3 (Słomnicka et al. 2018).

Within phylogroup 2, VCZ strains were mainly assigned to clade 2b-a and to clade 2a of the *P. syringae* species complex (Fig. 1B). Thirty-nine VCZ strains grouped in one highly supported cluster (bootstrap value of 100%) corresponding to clade 2b-a previously described by Newberry and his collaborators (2019). This cluster also gathered eight strains isolated from squash and watermelon (Newberry et al. 2019) and *P. syringae* pv. s*yringae* HS191 isolated from millet (*Panicum miliaceum*). The closest type and pathotype strains to clade 2b-a were *P. syringae* pv. *syringae* CFBP 1392 and *P. syringae* pv. *aptata* CFBP 1617 isolated from lilac (*Syringa vulgaris*) and sugar beet (*Beta vulgaris*), respectively. A second highly supported cluster (bootstrap value of 99%) previously

described as clade 2a (Berge et al. 2014), grouped 15 VCZ strains and four strains isolated from cantaloupe, squash, and watermelon (Newberry et al. 2019). The closest type and pathotype strains were *P. syringae* pv. *dysoxyli* CFBP 2356*, P. syringae pv. syringae* cit7 and *P. syringae* pv. *papulans* LMG 5076 isolated from kohekohe (*Dysoxylum spectabile*), orange (*Citrus sinensis*) and apple tree (*Malus domestica*), respectively. The last four VCZ strains ($P136_{2b\text{-}S}$, $P138_{2b\text{-}S}$, $P108_{2b\text{-}S}$ and $P101_{2b\text{-}S}$) grouped in clade 2b but were clearly separated from clade 2b-a. Three strains isolated from watermelon (Newberry et al. 2016) were closely related to these four singletons.

According to the SH test, ML trees performed on individual gene and the concatenated four genes were not congruent suggesting recombination events among loci (Supplementary fig. 1). Specifically, trees built on *gapA, gltA and rpoD* were not congruent with the tree based on the concatenated dataset. The differences were limited between *gltA* and the concatenated dataset trees, while clusters behaved substantially differently when *gapA* or *rpoD* were analyzed separately.

Pathogenicity of VCZ strains on zucchini. All VCZ strains, *P. amygdali* pv. *lachrymans* CFBP 6463 and *P. syringae* pv. *aptata* CFBP 5427 were pathogenic on zucchini as demonstrated with the cotyledon rubbing test, but only VCZ strains succeed to induce severe stunting (Fig. 2, Supplementary table 2). More specially, 23 VCZ strains were evaluated on two zucchini cultivars (*C. pepo* subsp. *pepo* cvs. Tosca and Opale). All VCZ strains induced cotyledon necrosis and leaf necrosis on plantlets with little variability among the tested strains (Fig. 2). Short-live vein clearing symptom was observed in the days following the inoculation of all VCZ strains, except for strains $P22_{2ba\text{-}A}$ and $P79_{2ba\text{-}A}$ that failed to induce vein clearing on one or the other cultivar (Supplementary table 2).

Most VCZ strains induced severe stunting except strains $P22_{2ba\text{-}A}$ and $P79_{2ba\text{-}A}$ which were less aggressive than other strains. Strains *P. syringae* pv. *aptata* CFBP 5427 and *P. amygdali* pv. *lachrymans* CFBP 6463 previously reported as pathogenic on various cucurbits (Bradbury 1986; Morris et al. 2000) were used as controls. Both strains induced necroses on both cultivars that were similar to those induced by VCZ strains. Moreover, *P. syringae* pv. *aptata* CFBP 5427 was able to produce short-live vein clearing on zucchini cv. Opale but not *P. amygdali* pv. *lachrymans* CFBP 6463. Neither CFBP 5427 nor CFBP 6463 were able to induce severe stunting on inoculated zucchini. No symptom was observed on plantlets inoculated with water and *P. syringae* pv. *dysoxyli* CFBP 2356, which is a pathotype strain from clade 2a that is not pathogenic on cucurbits.

Seed soaking tests revealed different levels of aggressiveness for VCZ strains and overall, no differences in terms of aggressiveness on zucchini were linked to the genetic group of the strains (Fig. 2). With the seed soaking test in a suspension at 1 x 10^6 ufc/mL, VCZ strains were highly aggressive on cv. Opale and weakly aggressive on cv. Tosca (Fig. 2). With the seed soaking test in a higher concentrated suspension (1 x 10^7 ufc/mL), VCZ strains were highly aggressive on both cultivars inducing necrosis on almost all plantlets. Strain $P22_{2ba\text{-}A}$ was the only VCZ strain that failed to induce any stunting with this test (Supplementary table 2). *P. amygdali* pv. *lachrymans* CFBP 6463 and *P. syringae* pv. *aptata* CFBP 5427 induced limited necroses on a few plantlets using these seed soaking tests. No symptoms were observed on plantlets issued from seeds inoculated with water and *P. syringae* pv. *dysoxyli* CFBP 2356.

Host range of VCZ strains. All VCZ strains induced similar symptoms on zucchini and pumpkin (*C. pepo* subsp. *pepo*) cultivars (Fig. 3A). On zucchini (*C. pepo* subsp. *pepo* cvs.

Tosca and Opale) and on pumpkin (*C. pepo* susp. *pepo* cv. Mischief), the strains responsible for VCZ P99$_{2ba-A}$, P79$_{2ba-A}$, P66$_{2a-E}$, P77$_{2a-D}$, P73$_{2a-E}$, P108$_{2b-S}$ induced water-soaking spots evolving in necrotic spots surrounded by a yellow halo at 7 dpi. Interestingly, *P. syringae* pv. *aptata* CFBP 5427 also induced this type of symptoms on these genotypes and could not be distinguished from the VCZ strains while *P. amygdali* pv. *lachrymans* CFBP 6463 induced smaller irregular necrotic lesions.

All VCZ strains were pathogenic on straightneck squash (*C. pepo* subsp. *texana*) and winter squashes (*C. maxima, C. moschata*) but different types of symptoms were observed according to the MLSA group (Fig. 3A). The yellow straightneck squash (*C. pepo* subsp. *texana* cv. Cheetah) seemed to be less susceptible to VCZ strains from clade 2b-a (P99$_{2ba-A}$, P79$_{2ba-A}$) which induced dried brown spots, while VCZ strains from clade 2a (P66$_{2a-E}$, P73$_{2a-D}$, P77$_{2a-D}$) induced water-soaked spots evolving in necrosis and were associated with chlorosis. On butternut (*C. moschata* cv. Sibelle) VCZ strains from clade 2b-a induced brown necrotic spots with a yellow halo that collapsed while strains from clade 2a induced large water-soaking lesions and necrotic spots surrounded by a yellow halo. On *C. maxima* cv. Rouge Vif d'Etampes, VCZ strains from clade 2b-a induced tiny yellow spots while VCZ strains from clade 2a induced water-soaked spots surrounded by a yellow halo.

Only VCZ strains from clade 2a and the singleton P108$_{2b-S}$ were pathogenic on the cucumber, melons and watermelon cultivars that were tested (*C. sativus, C. melo, C. lanatus*) (Fig. 3). VCZ strains from clade 2a (P66$_{2a-E}$, P73$_{2a-D}$, P77$_{2a-D}$) were pathogenic on all three melons varieties (*C. melo* var. *galia* cv. Citrex, *C. melo* var. *inodorus* cv. Helios, *C. melo* var. *cantaloupe* cv. Perseus) and induced necrotic spots surrounded by a yellow

halo. On cucumber (*C. sativus* cv. Dahab), VCZ strains from clade 2a were highly aggressive and induced collapsing circular necrotic spots with large chlorosis. In clade 2a, only strain $P66_{2a\text{-}E}$ was able to induce necrotic spots surrounded by a little yellow halo on watermelon (*C. lanatus* cv. Livia). The tested cultivars of cucumber, melon and watermelon were not susceptible to the VCZ strains from clade 2b-a ($P99_{2ba\text{-}A}$, $P79_{2ba\text{-}A}$). Some tiny dry white spots were sometimes visible on inoculated leaves but plants remained healthy two weeks post inoculation. Singleton VCZ strain $P108_{2b\text{-}S}$ was pathogenic on all tested cucurbits but the symptoms were different from those caused by VCZ strains from clade 2a ($P66_{2a\text{-}E}$, $P73_{2a\text{-}D}$, $P77_{2a\text{-}D}$) on some hosts. For instance, on watermelon, $P108_{2b\text{-}S}$ induced large dried necrotic spots while $P66_{2a\text{-}E}$ induced small spots with a discrete yellow halo. *P. amygdali* pv. *lachrymans* CFBP 6463 and *P. syringae* pv. *aptata* CFBP 5427 were pathogenic on almost all the tested genotypes, and the symptoms were sometimes different from those caused by VCZ strains. As expected, *P. amygdali* pv. *lachrymans* CFBP 6463 induced angular leaf spots on cucumber and *P. syringae* pv. *aptata* CFBP 5427 induced circular necrotic spots on melons. All plantlets inoculated with water remained asymptomatic 7 dpi.

Hypersensitive response (HR) of different cucurbits toward VCZ strains. Infiltration tests on cucumber, watermelon and melons confirmed that VCZ strains had different host range (Fig. 4). Starting from 48 hpi, spreading lesions that eventually collapsed with a water-soaked aspect on the abaxial leaf surface were noticed for all pathogenic controls on their hosts (*A. citrulli* on watermelon, *P. amygdali* pv. *lachrymans* on cucumber and *P. syringae* pv. *aptata* on melons). *P. syringae* pv. *dysoxyli* CFBP 2356, a non-pathogenic strain on cucurbits, induced HR with a dried brown necrosis limited to the infiltrated area.

As expected, strains P66$_{2a\text{-}E}$ and P108$_{2b\text{-}S}$ caused water-soaked lesions similar to that of the corresponding pathogenic controls on cucumber and the three melon varieties. In contrast, strain P99$_{2ba\text{-}A}$ caused the same HR-like necrosis than those of *P. syringae* pv. *dysoxyli* CFBP 2356 on all tested genotypes confirming non-host reaction. On watermelon, strain P66$_{2a\text{-}E}$ was pathogenic while strains P73$_{2a\text{-}D}$ and P77$_{2a\text{-}D}$ induced HR-like, confirming pathogenic heterogeneity on this host for strains of the clade 2a (Supplementary fig. 2).

VCZ strains genome sequence analyses. Based on whole genome sequences, ANIb revealed strong relationships between VCZ strains from 2a and 2b-a groups (Fig. 5, Supplementary table 3). First, ANIb values between genome sequences of *P. syringae* strains from clades 2a and 2b (except 2b-a strains) were below 95 % indicating that strains belonged to different species (Supplementary table 3). Interestingly, the genome sequences from the 23 VCZ strains and the other 14 cucurbit strains (Newberry et al. 2019) included in clades 2b-a and 2a were more similar with ANIb-value higher than 95 %.

Based on whole genome sequences, new clusters inside the main genetic groups (clades 2b-a and 2a) were highlighted by ANIb and shared *k*-mers analyses (Fig. 5). Both dendrograms displayed similar phylogeny and showed three new clusters within each clade 2b-a and 2a (Fig. 5). Within clade 2b-a, the first cluster (A) gathered 15 VCZ strains and two cucurbit strains (Ps711$_{2ba\text{-}A}$, 13-509A$_{2ba\text{-}A}$) isolated from squash (Newberry et al. 2019). The second cluster (B) gathered the strain P12857$_{2ba\text{-}B}$ responsible for VCZ with six cucurbit strains isolated from melon CC457 (Baltrus et al. 2014) and watermelon (13-C2$_{2ba\text{-}B}$, 13-140A$_{2ba\text{-}B}$, 14-410$_{2ba\text{-}B}$, 14-32$_{2ba\text{-}B}$, 14-Gil$_{2ba\text{-}B}$) and finally, cluster C did not

contain any VCZ strain but one strain (Zum3584$_{2ba-C}$) isolated from squash (Newberry et al. 2019). The strain *P. syringae* pv. *syringae* HS191 which was encompassed in clade 2b-a with MLSA was excluded from one of these clusters. Within clade 2a, the first cluster (D) comprised four VCZ strains (P77$_{2a-D}$, P73$_{2a-D}$, P78$_{2a-D}$, P87$_{2a-D}$) and two other strains (03-19A$_{2a-D}$, 200-1$_{2a-D}$) isolated from melon and squash (Newberry et al. 2019). The second cluster (E) gathered two VCZ strains (P66$_{2a-E}$, P89$_{2a-E}$) with strains Bs2121$_{2a-E}$ and Zum3984$_{2a-E}$ isolated from squash (Newberry et al. 2019). The last cluster (F) was composed of two strains (13-139B$_{2a-F}$, 13-149$_{2a-F}$) isolated from watermelon (Newberry et al. 2019).

Specific signatures were found in genome sequences with SkIf that allow to distinguish the six clusters containing cucurbit strains within clades 2a and 2b-a. In genomic sequences of strains from clade 2b-a, a total of 108 CDSs with specific signatures were found to differentiate the strains of cluster A from strains of clusters B and C. For cucurbit strains from clade 2a, a total of 578 CDSs with specific signatures of cluster E were found and allowed to distinguish cluster E from clusters D and F. Many specific *k*-mers were identified on CDSs localized in an integrative and conjugative element (ICE) but other signatures were located elsewhere in the genome sequences. Among these latter CDSs, three were retained because they were widely distributed within the set of 116 *P. syringae* genome sequences and allowed phylogenetic analyses. These three CDSs corresponded to *Psyr3420, Psyr4880* and *Psyr3208* in strain B728a of *P. syringae*, encoding the nitrogen fixation protein FixH, a hypothetical protein and the NADH dehydrogenase subunit M, respectively. OrthoMCL analysis revealed proteins specifically present in each cluster

comparatively to the others in the same clade. Therefore 13, 9, 190, 54, 75, and 384 proteins were specific of clusters A, B, C, D, E and F, respectively (data not shown).

Proposal of a novel MLSA scheme dedicated to cucurbit pathogens. Seven genes (*gapA, gltA, gyrB, rpoD, Psyr3420, Psyr4880* and *Psyr3208)* were selected to highlight all clusters observed with whole genome analyses within clades 2b-a and 2a. The MLSA tree based on the sequences of this novel MLSA scheme including these seven genes was constructed for 23 VCZ strains, 14 strains isolated from cucurbits (Newberry et al. 2019) and 22 diverse strains from phylogroup 2 (Fig. 6A). Phylogenies were constructed for each gene individually and for the concatenated data set (Supplementary fig. 3) and their topologies were compared with the SH test. As expected individual phylogenies were not congruent with the one constructed on the concatenated dataset indicating that each gene fragment was bringing complementary information. The tree showed that cucurbit strains of clade 2b-a were separated into three distinct branches and VCZ strains were carried by two of them. The cucurbit strains from clade 2a split also into three separate branches and VCZ strains were carried by two of them. Overall, composition of each branch was coherent with the composition of clusters identified using whole genome sequence analyses.

T3Es. Strains responsible for VCZ presented different T3E repertoires according to their genetic group (Fig. 7). Using t-BLASTN and sequence alignment, a total of 23 different T3Es were identified in the genome sequences of the 23 VCZ strains. The composition of T3E repertoires was associated with strain clusters identified with whole genome sequence analysis/core-genome MLSA as all strains from a given cluster had homogeneous set of T3Es (Fig. 6B, Fig. 7) except two strains from the collection of

Newberry, strains 14-Gil$_{2ba-B}$ and 13-139B$_{2a-F}$ which lacked one T3E. The 15 VCZ strains from the cluster A possessed a set of 18 T3Es and one of them (*avrRpt2*) was specific to that cluster. The strain P12857, which was the only VCZ strain from cluster B, possessed 16 T3Es in which 14 were common with the cluster A. The four VCZ strains from cluster D possessed 13 T3Es, of which one (*hopAR1*) was specific to strains from that cluster. The strains P89$_{2a-E}$ and P66$_{2a-E}$ from cluster E possessed 14 T3Es, 11 being common with cluster D. P108$_{2b-S}$ (singleton) had a smaller T3E repertoire (11 T3Es) with a unique composition but no specific T3E in comparison with the main VCZ clusters. Finally, eight T3Es (*avrE1, hopAA1, hopAE1, hopAZ1, hopAG1, hopAH1, hopBA1* and *hopI1)* were found in all strains responsible for VCZ.

The different genetic groups of VCZ strains differing by their host ranges are associated with different T3E repertoires (Fig. 6, Fig. 7). As no pathogenicity data was available for VCZ strains from cluster B, we evaluated the behavior of P12857$_{2ba-B}$ on cucumber, melons and watermelon with the infiltration method (Supplementary fig. 4). P12857$_{2ba-B}$ induced spreading water-soaked necrosis in all leaves from these hosts, thus indicating that this strain was pathogenic on these cucurbits. Interestingly, homogeneous T3E repertoires among strains from the same cluster gave homogeneous pathogenicity results (Fig. 6 and 7). The tested strains (P99$_{2ba-A}$, P79$_{2ba-A}$) from cluster A (clade 2b-a) were not pathogenic on cucumber, melons and watermelon and were weakly pathogenic on yellow straightneck genotype. Interestingly, strains from cluster A possessed a unique T3E (*avrRpt2*) and did not have any copy of *hopZ5* compared to all other VCZ strains from clusters B, D, E, which were highly pathogenic on these hosts. Strains P66$_{2a-E}$ (cluster E), P12857$_{2ba-B}$ (cluster B) and P108$_{2b-S}$ (singleton) were pathogenic on watermelon (*C.*

lanatus cv. Livia). It is worth noticing that all these strains lacked *hopC1* and *hopH1,* which were present in strains $P99_{2ba-A}$ and $P79_{2ba-A}$ from cluster A, and $P73_{2a-D}$ and $P77_{2a-D}$ from cluster D, strains that were not pathogenic and drove to a HR-like on watermelon.

Discussion

Facing the increase of VCZ in zucchini seed producing areas, we engaged a study to refine results obtained previously on this disease. The present study reveals that VCZ is currently caused by strains belonging to different phylogenetic lineages within the phylogroup 2 of the *P. syringae* species complex. Based on a MLSA using four housekeeping genes (*gapA, gltA, gyrB* and *rpoD*), VCZ strains are mainly grouped in two clades, namely clade 2b-a and clade 2a. A high-resolution analysis based on genome sequences of VCZ strains and cucurbit strains highlights six distinct clusters. VCZ strains belong to four of these six clusters: A and B being associated to clade 2b-a, while D and E are associated to clade 2a. Despite a common disease on zucchini (*C. pepo* subsp. *pepo*), VCZ strains induce different types of symptoms on other squashes (*C. pepo* subsp. *texana, C. moschata, C. maxima*) and have different host range of cucurbit genotypes. Genome sequences were mined to recover genes that matched those clusters. Three largely distributed genes (*Psyr3420, Psyr4880* and *Psyr3208)* were then selected to complete a novel MLSA scheme that can resolve clusters of strains pathogenic on cucurbits. The VCZ strains from each cluster share homogeneous T3E repertoires that are associated with pathogenicity features on cucurbit genotypes. Strains from cluster A have a narrow host range and are pathogenic on squashes (*Cucurbita* spp.), while VCZ strains from cluster B, D, E and singleton P108 have a wide host range including

528 squashes, cucumber and melons. By having different host range extents, VCZ strain
529 lineages represent different epidemiological risks in cucurbit producing areas.

530 Comparative genomic analyses led to proposing a novel MLSA scheme and differential
531 T3Es that distinguish various lineages of strains that are responsible for diseases on
532 cucurbits. VCZ strains group in four main lineages while a few strains remain isolated in
533 the phylogeny. Other strains responsible for diseases on cucurbits that were previously
534 characterized (Newberry et al. 2019) were included in our genomic analyses. Clearly, all
535 these strains are distant from *P. amygdali pv. lachrymans* and cause different diseases.
536 One lineage of VCZ strains (cluster A) presents a narrower host range than other groups,
537 including zucchini, yellow straightneck, butternut, and pumpkin. The presence of gene
538 encoding AvrRpt2 and the absence of gene encoding HopZ5 characterize this group of
539 VCZ strains. This lineage includes most strains previously named *P. syringae* pv. *peponis*
540 (Manceau et al. 2011) and is included in the clade 2b-a, which is described as being the
541 result of recombination events between *P. syringae* and *P. cerasi* (Newberry et al. 2019).
542 The three other groups of VCZ strains (clusters B, D and E) are distributed in clades 2a
543 and 2b-a. Clade 2a corresponds to *P. cerasi* (Newberry et al. 2019). We show here that
544 VCZ strains of these three groups are pathogenic not only on zucchini, but also on yellow
545 straightneck, butternut, pumpkins, cucumber, melons and in some cases watermelon.
546 These groups include also strains that were independently shown to induce angular leaf
547 spot and bacterial leaf spot of watermelon, cantaloupe and squash by Newbery and
548 colleagues (2016, 2018, 2019). The presence of the gene encoding HopZ5 and the
549 absence of the gene encoding AvrRpt2 in their genome sequences characterize these
550 lineages. Taking into account that, i) the diseases induced in zucchini by the four groups

551 of strains and the singleton ($P108_{2b\text{-}S}$ in clade 2b corresponding to *P. syringae sensu*
552 *stricto*) are similar, while the host range of cluster A is different from the one of clusters B,
553 D and E; ii) the distribution of these strains over different species and iii) the fact that we
554 did not test on our cucurbit genotypes the pathotype strain of. *P. syringae* pv. *syringae*,
555 known to have a large host range, we cannot conclude that the four lineages truly form
556 distinct pathovars. It should be noticed that the singleton $P108_{2b\text{-}S}$ is phylogenetically close
557 to *P. syringae* pv. *aptata* that includes strains isolated from melons (Berge et al. 2014).
558 Strain CFBP 5427 from this pathovar induces symptoms on squash, cucumber and
559 melons that are similar to those caused by $P108_{2b\text{-}S}$ strain. The only few differences are
560 the inability of the *P. syringae* pv. *aptata* CFBP 5427 strain to induce stunting of zucchini
561 plantlets with the seed soaking test and the susceptibility of watermelon to $P108_{2b\text{-}S}$ but
562 not to CFBP 5427.

563 Specific T3Es could be used as genetic markers of differential host ranges of VCZ strains.
564 In this study, we identified that presence of *hopZ5* and absence of *avrRpt2* are common
565 characters of strains from the three clusters sharing a wide host range, and having in
566 common to be pathogenic on melons and cucumber. In contrast cluster A strains have
567 *avrRpt2* and lack *hopZ5* and are not pathogenic on these groups of genotypes. Further
568 experimental work is necessary to decipher the role of these T3Es in host range extents.
569 Interestingly, it was also shown that strains from clusters C and cluster F harbor *hopZ5*
570 and are pathogenic on squash, melon and watermelon (Newberry et al. 2019). HopZ5 was
571 previously proposed to promote virulence among strains pathogenic on cucurbits and
572 having different T3E repertoires (Newberry et al. 2019). Here we suggest that *hopZ5* could
573 also be a good predictor of a wide host range on cucurbits. Cluster A strains ($Ps711_{2ba\text{-}A}$

and 13-509A$_{2ba-A}$), harboring *avrRpt2* and missing *hopZ5* have been previously reported to be weakly pathogenic on squash, melon and watermelon (Newberry et al. 2018, 2019). Their pathogenicity was evaluated on a yellow straightneck genotype but strains were not tested on cucumber nor on any other genotypes of squashes, including zucchini. Therefore, this weak pathogenicity on yellow straightneck for cluster A strains is coherent with the dried brown lesions we observed for strains P99$_{2ba-A}$ and P79$_{2ba-A}$ on yellow straightneck. Indeed, the dried lesions seemed less severe than the water-soaked lesions caused by strains of clusters B, D and E and the singleton. The gene *avrRpt2* is commonly found in *P. syringae* pv. *tomato* strains (Almeida et al. 2008) and has been shown to promote pathogenicity by stimulating key regulators in auxin signaling pathways (Cui et al. 2013). This effector leads to HR-like in *A. thaliana* through the cognate RPS2 resistance gene (Pike et al. 2005). Apart from the main genetic clusters including VCZ strains, we also characterized a singleton (P108$_{2b-S}$), which is pathogenic on all tested cucurbits and induces VCZ. P108$_{2b-S}$ is close to *P. syringae* pv. *aptata* strains in the phylogenetic analysis. This singleton has fewer T3Es (11 T3Es) than other VCZ strains. Seven T3Es (*avrE1, hopAA1, hopAE1, hopAZ1, hopAG1, hopAH1, hopBA1* and *hopI1*) are common to all the VCZ strains, and might be associated to pathogenicity on zucchini. It is also interesting to note that Newberry and collaborators characterized singletons from clade 2b and 2d isolated from diseased cucurbits (Newberry et al. 2016, 2019). Previous studies also identified a potential large host range for strains from clade 2d (Monteil et al. 2016; Morris et al. 2019).

Not all VCZ strains were pathogenic on watermelon cultivar Livia as tested strains from clusters A (P99$_{2ba-A}$ and P79$_{2ba-A}$) and D (P73$_{2a-D}$ and P77$_{2a-D}$) did not induce any disease

with the spray inoculation and induced HR-like necrosis when infiltrated in leaves. These results were also obtained with cv. Troubadour (data not shown). Based on genome analysis, strains Ps711$_{2ba-A}$ and 13-509A$_{2ba-A}$ belong to cluster A and strains 200-1$_{2a-D}$ and 03-19A$_{2a-D}$ belong to cluster D. Interestingly, these strains were previously tested with a similar spray inoculation and were shown to be weakly pathogenic or not pathogenic on watermelon cultivars Troubadour and Wrigley (Fig. 6B) (Newberry et al. 2016, 2017, 2019). These discrepancies on watermelon could be strain- or test-dependent. However, some effectors are associated to strains from cluster A and D which are weakly pathogenic or not pathogenic on watermelon as all possess complete genes of *hopH1* and *hopC1* in contrast to other aggressive strains from cluster B, C, E, F and singleton. This result is coherent with those previously reported (Newberry et al. 2019).

The use of several pathogenicity tests permitted a precise phenotypic characterization of strains responsible for VCZ. On zucchini, the cotyledon rubbing test successfully reproduces the typical symptoms (necroses on leaves and cotyledons, vein clearing and stunting) observed in seedling producing areas and caused by VCZ strains. The cotyledon rubbing test is quick and easy to perform, however it induces micro-injuries on plantlets by trichome breaks and thus creates openings. For this reason, the cotyledon rubbing test is quite invasive and can produce biases to study bacterial strains pathogenicity (Burdman et al. 2005; Klement et al. 1990). The seed soaking tests were appropriate to study VCZ strains on their host of origin and permitted to properly characterize strain aggressiveness. Seed soaking test is coherent with VCZ strains being described as seed-borne bacteria (Manceau et al. 2011). Seed soaking tests are commonly used to characterize aggressiveness of seed-borne bacterium such as *A. avenae* subsp. *citrulli* on cucurbits

(Burdman et al. 2005). Then, the spray inoculation method was used to simulate natural infection of leaves by pathogen and allowed bacterial entry through natural openings (Klement et al. 1990). This test was particularly relevant to finely characterize symptoms and host range. Finally, the HR test is an easy, fast and efficient method to define host range of VCZ strains on different cucurbit genotypes. Host-incompatibility gives HR-like reaction on leaves after 48 h post-inoculation while host-compatibility gives spreading water-soaked necrosis.

The strains responsible for VCZ were isolated from plantlets issued from zucchini seed lots produced in geographic distant regions over time. Only few studies so far deal with *P. syringae* strains transmitted by cucurbits seeds (Dutta et al. 2014; Meng et al. 2016; Shila et al. 2013). Strains from *P. syringae* pv. *aptata*, *P. syringae* pv. *syringae* and *P. amygdali* pv. *lachrymans* had been reported to be seed disseminated and efforts to develop seed detection tests contribute to improve pathogen control and disease management. We identified specific markers (*Psyr3420, Psyr4880 and Psyr3208*) that can be used to design specific VCZ strain primers. A specific detection tool is indeed needed to conduct epidemiological studies of VCZ in worldwide zucchini seed production areas. Typing *P. syringae* isolates from diseased cucurbits with the new MLSA scheme would give insight into worldwide dissemination of strains and origin of epidemics. The gain of knowledge on VCZ strains and the survey of epidemics are essential steps to implement an efficient management of the various VCZ lineages, which do not represent the same risk as a consequence of their different host range.

Acknowledgments

643 We thank P. Portier and C. Dutrieux (CIRM-CFBP) for strain preservation. We are also
644 grateful to the SFR Quasav and the ANAN platform, in Angers, France, for sequencing
645 facilities. The authors thank M. Barret, A. Sarniguet, M. Simonin and three anonymous
646 reviewers for their very fruitful reviews of the manuscript.

Almeida, N. F., Yan, S., Lindeberg, M., Studholme, D. J., Schneider, D. J., Condon, B., et al. 2008. A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. Mol. Plant. Microbe Interact. 22:52–62.

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

Baltrus, D. A., Yourstone, S., Lind, A., Guilbaud, C., Sands, D. C., Jones, C. D., et al. 2014. Draft genome sequences of a phylogenetically diverse suite of *Pseudomonas syringae* strains from multiple source populations. Genome Announc. 2(1):01195-13.

Baltrus, D., Honour, C., and Guttman, S. 2017. Evolution, genomics and epidemiology of *Pseudomonas syringae*. Mol. Plant Pathol. :152–168.

Berge, O., Monteil, C. L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D. C., et al. 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(9):e105547.

Bradbury, J. F. 1986. Guide to plant pathogenic bacteria. CAB International Mycological Institute.

Briand, M., Bouzid, M., Hunault, G., Legeay, M., Saux, M. F.-L., and Barret, M. 2019. A rapid and simple method for assessing and representing genome sequence relatedness. bioRxiv. :569640.

Bull, C. T., Boer, S. H. D., Denny, T. P., Firrao, G., Saux, M. F.-L., Saddler, G. S., et al. 2008. Demystifying the nomenclature of bacterial plant pathogens. J. Plant Pathol. 90:403–417.

Bull, C. T., Clarke, C. R., Cai, R., Vinatzer, B. A., Jardini, T. M., and Koike, S. T. 2011. Multilocus sequence typing of *Pseudomonas syringae* sensu lato confirms previously described genomospecies and permits rapid identification of *P. syringae* pv. *coriandricola* and *P. syringae* pv. *apii* causing bacterial leaf spot on parsley. Phytopathology 101:847-858.

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.

Burdman, S., Kots, N., Kritzman, G., and Kopelowitz, J. 2005. Molecular, physiological, and host-range characterization of *Acidovorax avenae* subsp. *citrulli* isolates from watermelon and melon in israel. Plant Dis. 89:1339–1347.

Corpet, F., Gouzy, J., and Kahn, D. 1999. Browsing protein families via the 'Rich Family Description' format. Bioinformatics 15:1020-1027.

Cottret, L., Rancurel, C., Briand, M., and Carrere, S. 2018. Family-Companion: analyse, visualise, browse, query and share your homology clusters. Bioinformatics.

Cui, F., Wu, S., Sun, W., Coaker, G., Kunkel, B., He, P., et al. 2013. The *Pseudomonas syringae* type III effector *avrRpt2* promotes pathogen virulence via stimulating *Arabidopsis* auxin/indole acetic acid protein turnover. Plant Physiol. 162:1018–1029.

Cunty, A., Poliakoff, F., Rivoal, C., Cesbron, S., Fischer-Le Saux, M., Lemaire, C., et al. 2015. Characterization of *Pseudomonas syringae* pv. *actinidiae* (Psa) isolated from France and assignment of Psa biovar 4 to a de novo pathovar: *Pseudomonas syringae* pv. *actinidifoliorum* pv. nov. Plant Pathol. 64:582–596.

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

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

Dillon, M. M., Almeida, R. N. D., Laflamme, B., Martel, A., Weir, B. S., Desveaux, D., et al. 2019b. Molecular Evolution of *Pseudomonas syringae* Type III Secreted Effector Proteins. Front. Plant Sci. 10 :418.

Dutta, B., Ingram, T., Gitaitis, R. D., Langston, D. B., Brenneman, T., Webster, T. M., et al. 2014. First report of bacterial blight of sugar beet caused by *Pseudomonas syringae* pv. *aptata* in Georgia, USA. Plant Dis. 98:1423–1423.

Dutta, B., Gitaitis, R., Agarwal, G., Coutinho, T., and Langston, D. 2018. *Pseudomonas coronafaciens* sp. nov., a new phytobacterial species diverse from *Pseudomonas syringae*. PLoS ONE 13: e0208271.

Dye, D. W., Bradbury, J. F., Goto, M., Hayward, A. C., Lelliott, R. A., and Schroth, M. N. 1980. International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. Rev. Plant Pathol. 59:153–168.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics. 5:164–166.

Feng, F., and Zhou, J.-M. 2012. Plant-bacterial pathogen interactions mediated by type III effectors. Curr. Opin. Plant Biol. 15:469–476.

715 Ferrante, P., and Scortichini, M. 2015. Redefining the global populations of *Pseudomonas*
716 *syringae* pv. *actinidiae* based on pathogenic, molecular and phenotypic characteristics.
717 Plant Pathol. 64:51–62.
718 Gomila, M., Busquets, A., Mulet, M., García-Valdés, E., and Lalucat, J. 2017. Clarification
719 of taxonomic status within the *Pseudomonas syringae* species group based on a
720 phylogenomic analysis. Front. Microbiol. 8:24-22.
721 Hopkins, D., and Schenck, N. 1972. Bacterial leaf spot of watermelon caused by
722 *Pseudomonas-lachrymans*. Phytopathology. 62:542.
723 Hwang, M. S. H., Morgan, R. L., Sarkar, S. F., Wang, P. W., and Guttman, D. S. 2005.
724 Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas*
725 *syringae*. Appl. Environ. Microbiol. 71:5182–5191.
726 Jacques, M.-A., Durand, K., Orgeur, G., Balidas, S., Fricot, C., Bonneau, S., et al. 2012.
727 Phylogenetic analysis and polyphasic characterization of *Clavibacter michiganensis*
728 strains isolated from tomato seeds reveal that nonpathogenic strains are distinct from *C.*
729 *michiganensis* subsp. *michiganensis*. Appl. Environ. Microbiol. 78:8388–8402.
730 Jacques, M.-A., Denance, N., Legendre, B., Morel, E., Briand, M., Mississipi, S., et al.
731 2016. New coffee plant-infecting *Xylella fastidiosa* variants derived via homologous
732 recombination. Appl. Environ. Microbiol. 82:1556–1568.
733 Kaluzna, M., Willems, A., Pothier, J.F., Ruinelli, M., Sobiczewski, P., and Pulawska, J.
734 2016. *Pseudomonas cerasi* sp nov (non Griffin, 1911) isolated from diseased tissue of
735 cherry. Syst. Appl. Microbiol. 39: 370-77.
736 Klement, Z., Rudolph, K., and Sands, D. C. 1990. Methods in phytobacteriology.
737 Akadémiai Kiado, Budapest, Hungary.

738 Konstantinidis, K. T., and Tiedje, J. M. 2005. Towards a genome-based taxonomy for
739 prokaryotes. J. Bacteriol. 187:6258–6264.
740 Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular Evolutionary Genetics
741 Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33:1870–1874.
742 Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. 2012. SOAPdenovo2: an
743 empirically improved memory-efficient short-read de novo assembler. GigaScience. 1:18.
744 Manceau, C., Gironde, S., Briand, B., and Lybeert, H. 2011. Means and methods for
745 detecting and identifying a novel bacterium responsible for phytosanitary disorders in
746 plants (zucchini) and novel resistant plants. Patent #WO2011003984. World Intellectual
747 Property Organization, Geneva, Switzerland.
748 Martin-Sanz, A., Perez de la Vega, M., Murillo, J., and Caminero, C. 2013. Strains of
749 *Pseudomonas syringae* pv. *syringae* from pea are phylogenetically and pathogenically
750 diverse. Phytopathology. 103:673–681.
751 Meng, X., Chai, A., Chen, L., Shi, Y., Xie, X., Ma, Z., et al. 2016. Rapid detection and
752 quantification of viable *Pseudomonas syringae* pv. *lachrymans* cells in contaminated
753 cucumber seeds using propidium monoazide and a real-time PCR assay. Can. J. Plant
754 Pathol. 38:296–306.
755 Monteil, C. L., Cai, R., Liu, H., Llontop, M. E., Leman, S., Studholme, D. J., Morris, C. E.,
756 and Vinatzer, B. A. 2013. Nonagricultural reservoirs contribute to emergence and
757 evolution of *Pseudomonas syringae* crop pathogens. New Phytol. 199:800-811.
758 Monteil, C. L., Yahara, K., Studholme, D. J., Mageiros, L., Méric, G., Swingle, B., et al.
759 2016. Population-genomic insights into emergence, crop adaptation and dissemination of
760 *Pseudomonas syringae* pathogens. Microb. Genom. 21:e000089.

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., Lamichhane, J. R., Nikolic, I., Stanković, S., and Moury, B. 2019. The overlapping continuum of host range among strains in the *Pseudomonas syringae* complex. Phytopathol. Res. 1:4

Newberry, E. A., Jardini, T. M., Rubio, I., Roberts, P. D., Babu, B., Koike, S. T., et al. 2016. Angular leaf spot of cucurbits is associated with genetically diverse *Pseudomonas syringae* strains. Plant Dis. 100:1397–1404.

Newberry, E. A., Ritchie, L., Babu, B., Sanchez, T., Beckham, K. A., Jones, J. B., et al. 2017. Epidemiology and management of bacterial leaf spot on watermelon caused by *Pseudomonas syringae*. Plant Dis. 101:1222–1229.

Newberry, E. A., Babu, B., Roberts, P. D., Dufault, N. S., Goss, E. M., Jones, J. B., et al. 2018. Molecular epidemiology of *Pseudomonas syringae* pv. *syringae* causing bacterial leaf spot of watermelon and squash in Florida. Plant Dis. 102:511–518.

Newberry, E. A., Ebrahim, M., Timilsina, S., Zlatković, N., Obradović, A., Bull, C. T., et al. 2019. Inference of convergent gene acquisition among *Pseudomonas syringae* strains isolated from watermelon, cantaloupe, and squash. Front. Microbiol. 10:270.

Paris, H. S. 1989. Historical records, origins, and development of the edible cultivar groups of *Cucurbita pepo* (*Cucurbitaceae*). Econ. Bot. 43:423–443.

Paris, H. S., Doron-Faigenboim, A., Reddy, U. K., Donahoo, R., and Levi, A. 2015. Genetic relationships in *Cucurbita pepo* (pumpkin, squash, gourd) as viewed with high frequency

784 oligonucleotide-targeting active gene (HFO-TAG) markers. Genet. Resour. Crop Evol.
785 62:1095–1111.
786 Pike, S. M., Zhang, X.-C., and Gassmann, W. 2005. Electrophysiological characterization
787 of the *Arabidopsis avrRpt2*-specific hypersensitive response in the absence of other
788 bacterial signals. Plant Physiol. 138:1009–1017.
789 Sallet, E., Gouzy, J., and Schiex, T. 2014. EuGene-PP: a next-generation automated
790 annotation pipeline for prokaryotic genomes. Bioinformatics. 30:2659–2661.
791 Sarkar, S. F., and Guttman, D. S. 2004. Evolution of the core genome of *Pseudomonas*
792 *syringae*, a highly clonal, endemic plant pathogen. Appl. Environ. Microbiol. 70:1999–
793 2012.
794 Sedighian, N., Shams-Bakhsh, M., Osdaghi, E., and Khodaygan, P. 2014. Etiology and
795 host range of bacterial leaf blight and necrosis of squash and muskmelon in Iran. J. Plant
796 Pathol. 96: 507-514.
797 Shila, S. J., Islam, M. R., Ahmed, N. N., Dastogeer, K. M. G., and Meah, M. B. 2013.
798 Detection of *Pseudomonas syringae* pv. *lachrymans* associated with the seeds of
799 cucurbits. Univers. J. Agric. Res. 1:1–8.
800 Shimodaira, H., and Hasegawa, M. 1999. Multiple Comparisons of Log-Likelihoods with
801 Applications to Phylogenetic Inference. Mol. Biol. Evol. 16:1114–1114.
802 Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., and Jackson, R. W. 2011.
803 *Pseudomonas* genomes: diverse and adaptable. Fems Microbiol. Rev. 35:652–680.
804 Słomnicka, R., Olczak-Woltman, H., Bartoszewski, G., and Niemirowicz-Szczytt, K. 2015.
805 Genetic and pathogenic diversity of *Pseudomonas syringae* strains isolated from
806 cucurbits. Eur. J. Plant Pathol. 141:1–14.

Słomnicka, R., Olczak-Woltman, H., Oskiera, M., Schollenberger, M., Niemirowicz-Szczytt, K., and Bartoszewski, G. 2018. Genome analysis of *Pseudomonas syringae* pv. *lachrymans* strain 814/98 indicates diversity within the pathovar. Eur. J. Plant Pathol. 151:663–676.

Tian, Y., Zhao, Y., Chen, X., Dai, Y., Zhao, W., Hu, B., et al. 2017. Evidence for a novel phylotype of *Pseudomonas syringae* causing bacterial leaf blight of cantaloupe in China. Plant Dis. 101:1746–1752.

Tymon, L. S., and Inglis, D. A. 2017. Identification and pathogenicity of *Pseudomonas syringae* genomospecies 1 phylogroup 2b causing leaf spots and fruit warts on cucurbits in western washington, U.S. J. Plant Pathol. 99:713–722.

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 (1, Supplement), S1.5-S1.14.

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

TABLE 1. 58 VCZ strains used in this study

VCZ strain number	Origin	Year	Host	Genome accession
P1	France	2005	zucchini	
P100	na	2011	zucchini	
P101	na	2011	zucchini	
P102	na	2011	zucchini	
P104	na	2011	zucchini	
P108	Chile	2011	zucchini	VLIA00000000
P112	Chile	2011	zucchini	
P113	Thailand	2011	zucchini	VLIC00000000
P115	China	2011	zucchini	
P118	France	2013	zucchini	VLID00000000
P119	France	2013	zucchini	
P12	France	2006	zucchini	
P120	France	2013	zucchini	
P121	Chile	2013	zucchini	VLIE00000000
P123	China	2013	zucchini	VLIF00000000
P125	China	2013	zucchini	
P127	India	2013	zucchini	VLIG00000000
P12831	France	2005	zucchini	WJPZ00000000
P12832	France	2006	zucchini	VLHX00000000
P12833	France	2005	zucchini	
P12836	France	2005	zucchini	
P12854	France	2002	zucchini	VLHY00000000
P12855	Chile	2007	zucchini	VLHZ00000000
P12856	France	2005	zucchini	
P12857	France	2007	zucchini	VLIB00000000
P130	Thailand	2013	zucchini	
P131	China	2013	zucchini	
P133	Thailand	2013	zucchini	
P136	India	2014	zucchini	
P137	France	2014	zucchini	
P138	France	2015	zucchini	
P139	China	2015	zucchini	VLIH00000000
P14	na	2008	zucchini	
P22	USA	2009	zucchini	WJPQ00000000
P24	France	2008	zucchini	
P4	France	2005	zucchini	
P5	France	2008	zucchini	WJPR00000000
P55	Chile	2009	zucchini	
P56	na	2009	zucchini	
P62	France	2010	zucchini	
P65	na	2010	zucchini	
P66 (CFBP 8692)	Thailand	2010	zucchini	WJSF00000000
P68	USA	2010	zucchini	
P69	USA	2010	zucchini	
P70	Thailand	2010	zucchini	
P71	na	2010	zucchini	
P72	na	2010	zucchini	
P73	USA	2010	zucchini	WJPS00000000
P77	France	2010	zucchini	WJPT00000000
P78	France	2010	zucchini	WJPU00000000
P79	France	2010	zucchini	WJPV00000000
P84	China	2010	zucchini	WJPW00000000
P87	China	2011	zucchini	WJPX00000000
P88	China	2011	zucchini	
P89	France	2011	zucchini	WJPY00000000
P90	USA	2011	zucchini	
P94	na	2011	zucchini	
P99 (CFBP 8693)	France	2005	zucchini	WJSE00000000

Fig. 1. Phylogenetic relationships of strains responsible for VCZ. The diversity of 58 VCZ strains and 17 strains isolated from diverse cucurbits species (Newberry et al. 2016, 2019) was assessed through MLSA. Maximum likelihood (ML) trees were based on concatenated partial sequences of *gapA*, *gltA*, *gyrB* and *rpoD* (2007 bp). Bootstrap scores (1,000 replicates) are displayed at each node. (A) ML tree of typed strains representing the *P. syringae* group. Phylogroups (PGs) are derived from Bull and Koike (2015) (B) Detail of the phylogroup 2 diversity including VCZ strains. Clades are indicated based on Berge et al. 2014; Bull and Koike 2015; and Newberry et al. 2019. Black points indicate VCZ strains.

Fig. 2. Responses of two cultivars of zucchini to seed inoculation and cotyledon rubbing assay with *P. syringae* strains. *Cucurbita pepo* subsp. pepo cv. *Opale* (orange) and cv. *Tosca* (blue) plants or seeds were inoculated with 23 strains of *P. syringae* responsible for VCZ and four control strains including CFBP 5427 (*P. syringae* pv. *aptata*), CFBP 6463 (*P. amygdali* pv. *lachrymans*), CFBP 2356 (*P. syringae* pv. *dysoxyli*) and water. Bars indicate the percentage of disease incidence over the number of germinated plantlets seven days post-inoculation for a given strain and a given test (cotyledon rubbing, number of tested plantlets n = 5; seed soaking at 1x 10^7 CFU/mL, number of sown seeds n = 10; seed soaking at 1 x 10^6 CFU/mL, number of sown seeds n = 10). (*) SS for seed soaking test.

Fig. 3. Host range test of VCZ strains. Six *P. syringae* strains responsible for VCZ and three controls including water, *P. syringae* pv. *aptata* CFBP 5427 and *P. amygdali* pv. *lachrymans* CFBP 6463 were spray inoculated (1 x 10^7 cfu/ml until runoff) on six genotypes of *Cucurbita* sp., four genotypes of *Cucumis* sp. and one genotype of *Citrullus lanatus*. Water was sprayed as control. (A) Photography of the symptoms observed on the first leaf 7 dpi (B) Results of pathogenicity test, (+) disease symptoms, (-) absence of symptoms.

Fig. 4. Hypersensitive response of diverse cucurbits following infiltration of VCZ strains. Three VCZ strains (P99, P66 and P108) were infiltrated in leaves of four genotypes of *Cucumis*

sp. and one *Citrullus lanatus*. Symptoms were pictured 72 h post-inoculation. Pathogenic controls consisted in *A. citrulli* CFBP 4459 on *C. lanatus* cv. *Livia*, *P. syringae* pv. *aptata* CFBP 5467 on melon cultivars and *P. amygdali* pv. *lachrymans* CFBP6463 on cucumber. The HR control for all tested plants was *P. syringae* pv. *dysoxily* CFBP 2356. Water was used as a negative control. Within the inoculated area the red line delimits necrosis and the orange line delimits water-soaked lesion. The white line delimits the expansion of water-soaked symptoms beyond the inoculation point.

Fig. 5. Relationships between genomic sequences of *P. syringae* from phylogroup 2 using ANIb and k-mers approaches. Dendrograms were built on distances calculated from ANIb and k-mers matrices obtained with Ki-S (Briand et al. 2019). The color blocks represent the cucurbit clusters highlighted with both approaches.

Fig. 6. Phylogeny of VCZ strains, host range and presence of some T3E encoding genes. (A) Phylogeny based on new MLSA scheme using seven genes (*gapA, gltA, gyrB, rpoD, Psyr3420, Psyr4880* and *Psyr3208).* Maximum likelihood tree was based on concatenated partial sequences (4431 pb). Bootstrap scores (1,000 replicates) were displayed at each node. Black points indicate VCZ strains. (B) Summarized results obtained from pathogenicity tests and presence of differential genes encoding T3Es recovered from genome sequence analysis. Columns in dark grey report results from Newberry et al. 2018, 2019 that were available for strains Ps711, 13-509A, 13-140A, 13-C2, 14-32, 14-410, 14-Gil, Zum3584, 13-139B, 13-429, Zum3984, Bs2121, 03-19A and 200-1. The light grey report the results obtained in this study. Colors are given as following: light or dark grey when a strain is not tested on a given host, red when a strain is pathogenic on a given host, yellow when a strain is weakly pathogenic on a given host and green when a strain is not pathogenic on a given host. A black box means a T3E presence. Cucurbit species and sub-species are reported as following, ZU for zucchini, YS for yellow Straightneck squash, PU for pumpkin, ME for melons and WM for watermelon.

Fig. 7. Set of T3Es predicted in the genome sequences of VCZ strains and other strains isolated from cucurbits Color blocks reflect presence of the effector within a cluster (yellow: cluster A, orange: cluster B, blue: cluster C, green: cluster E, purple: cluster D, grey: cluster F, pink: singleton P108). White block means the T3E absence.

Supplementary table 1. List of strains and their characteristics and list of genome sequences used in this study.

Supplementary table 2. Plantlets evaluation for the cotyledon rubbing test and seed soaking test at 1.10^7 cfu/mL. Colors refer to the incidence of a given symptom, with red for more than 80% of plantlets, orange 40 to 79%, yellow for 1 to 39% and green when none of the plantlets presented the symptom. For vein clearing, red color indicate that this symptom was observed at least at one date and green color when it had been never observed.

Supplementary table 3. ANIb-values calculated on genomic sequences.

Supplementary fig. 1. Phylogenies built on four individual housekeeping genes (*gapA, gltA, gyrB* and *rpoD*) and the concatenated dataset. Maximum-likelihood (ML) trees were generated with MEGA 7 (Kumar et al. 2016) using the Tamura-Nei model with gamma distributed with invariant sites (G+I), and 1000 bootstrap replicates. Blue color represents the clade 2b-a and red color the cucurbit clade 2a. P values determined using the Shimodaira-Hasegawa test for congruence of tree topologies run on each tree based on four housekeeping genes, individually and the concatenated dataset. Values an asterisk are significantly depart from the hypothesis of congruency.

Supplementary fig. 2. Response of watermelon (*C. lanatus* cv. Livia) to infiltration of *P. syringae* VCZ strains P66, P73, P77 and P79.

Supplementary fig. 3. Phylogenies built on seven individual genes (*gapA, gltA, gyrB*, *rpoD, Psyr3208, Psyr3420, Psyr4880*) and the concatenated sequences. Maximum-likelihood (ML) trees

were generated with MEGA 7 (Kumar et al. 2016) using the Tamura-Nei model with gamma distributed with invariant sites (G+I), and 1000 bootstrap replicates. P values determined using the Shimodaira-Hasegawa test for congruence of tree topologies run on each tree based on four housekeeping genes, individually and the concatenated dataset. Values an asterisk are significantly depart from the hypothesis of congruency.

Supplementary fig. 4. Response of melons, cucumber and watermelon to infiltration of VCZ strain P12857 from clade 2b-a.

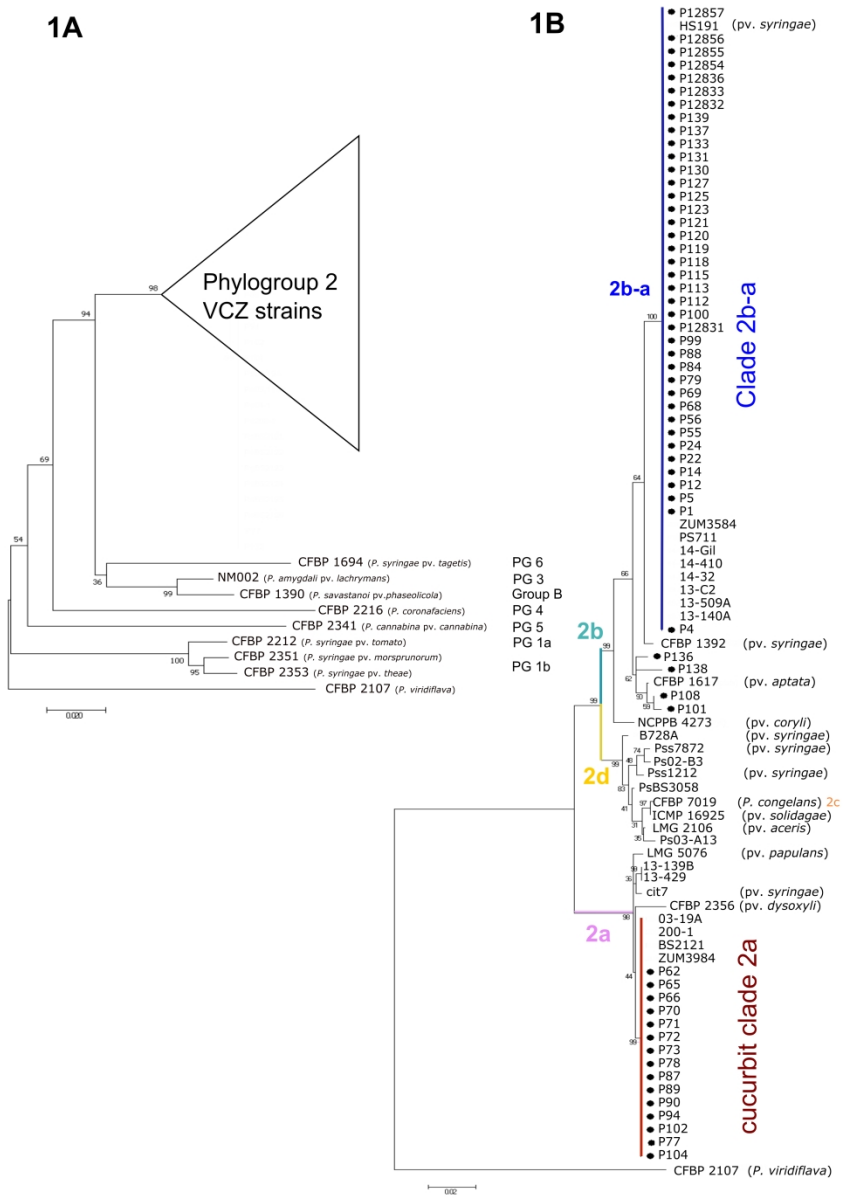


Fig. 1. Phylogenetic relationships of strains responsible for VCZ. The diversity of 58 VCZ strains and 17 strains isolated from diverse cucurbits species (Newberry et al. 2016, 2019) was assessed through MLSA. Maximum likelihood (ML) trees were based on concatenated partial sequences of gapA, gltA, gyrB and rpoD (2007 bp). Bootstrap scores (1,000 replicates) are displayed at each node. (A) ML tree of typed strains representing the P. syringae group. Phylogroups (PGs) are derived from Bull and Koike (2015) (B) Detail of the phylogroup 2 diversity including VCZ strains. Clades are indicated based on Berge et al. 2014; Bull and Koike 2015; and Newberry et al. 2019. Black points indicate VCZ strains.

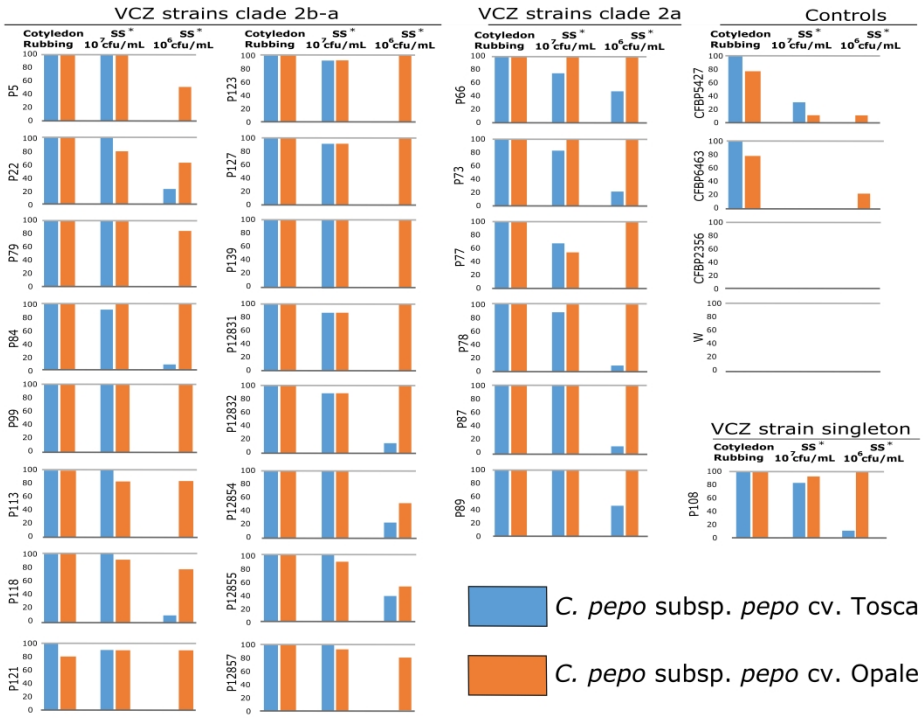
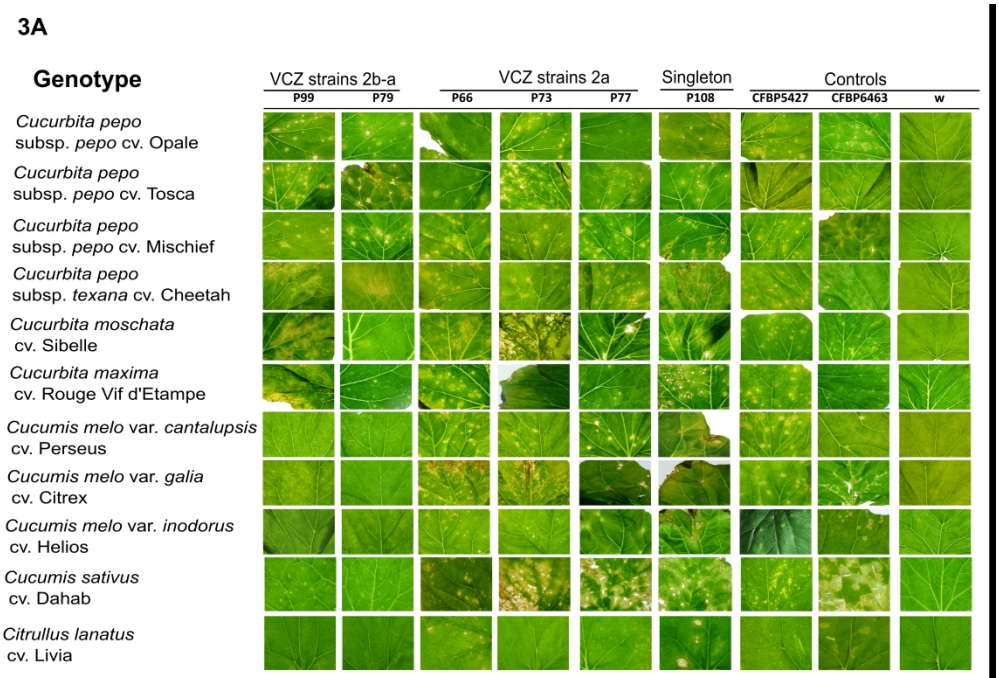


Fig. 2. Responses of two cultivars of zucchini to seed inoculation and cotyledon rubbing assay with P. syringae strains. Cucurbita pepo subsp. pepo cv. Opale (orange) and cv. Tosca (blue) plants or seeds were inoculated with 23 strains of P. syringae responsible for VCZ and four control strains including CFBP 5427 (P. syringae pv. aptata), CFBP 6463 (P. amygdali pv. lachrymans), CFBP 2356 (P. syringae pv. dysoxyli) and water. Bars indicate the percentage of disease incidence over the number of germinated plantlets seven days post-inoculation for a given strain and a given test (cotyledon rubbing, number of tested plantlets n = 5; seed soaking at 1x 107 CFU/mL, number of sown seeds n = 10; seed soaking at 1 x 106 CFU/mL, number of sown seeds n = 10). (*) SS for seed soaking test.

3A



3B

Genus	Genotype	VCZ strains 2b-a		VCZ strains 2a			Singleton	Controls		
		P99	P79	P66	P73	P77	P108	CFBP5427	CFBP6463	W
Curcurbita	*Cucurbita pepo* subsp. *pepo* cv. Opale	+	+	+	+	+	+	+	+	-
	Cucurbita pepo subsp. *pepo* cv. Tosca	+	+	+	+	+	+	+	+	-
	Cucurbita pepo subsp. *pepo* cv. Mischief	+	+	+	+	+	+	+	+	-
	Cucurbita pepo subsp. *texana* cv. Cheetah	+	+	+	+	+	+	+	+	-
	Cucurbita moschata cv. Sibelle	+	+	+	+	+	+	+	+	-
	Cucurbita maxima cv. Rouge vif d'Etampe	+	+	+	+	+	+	+	+	-
Cucumis	*Cucumis melo* var. *cantalupsis* cv. Perseus	-	-	+	+	+	+	+	+	-
	Cucumis melo var. *galia* cv. Citrex	-	-	+	+	+	+	+	+	-
	Cucumis melo var. *inodorus* cv. Helios	-	-	+	+	+	+	+	+	-
	Cucumis sativus cv. Dahab	-	-	+	+	+	+	+	+	-
Citrullus	*Citrullus lanatus* cv. Livia	-	-	+	-	-	+	-	+	-

Fig. 3. Host range test of VCZ strains. Six P. syringae strains responsible for VCZ and three controls including water, P. syringae pv. aptata CFBP 5427 and P. amygdali pv. lachrymans CFBP 6463 were spray inoculated (1 x 107 cfu/ml until runoff) on six genotypes of Cucurbita sp., four genotypes of Cucumis sp. and one genotype of Citrullus lanatus. Water was sprayed as control. (A) Photography of the symptoms observed on the first leaf 7 dpi (B) Results of pathogenicity test, (+) disease symptoms, (-) absence of symptoms.

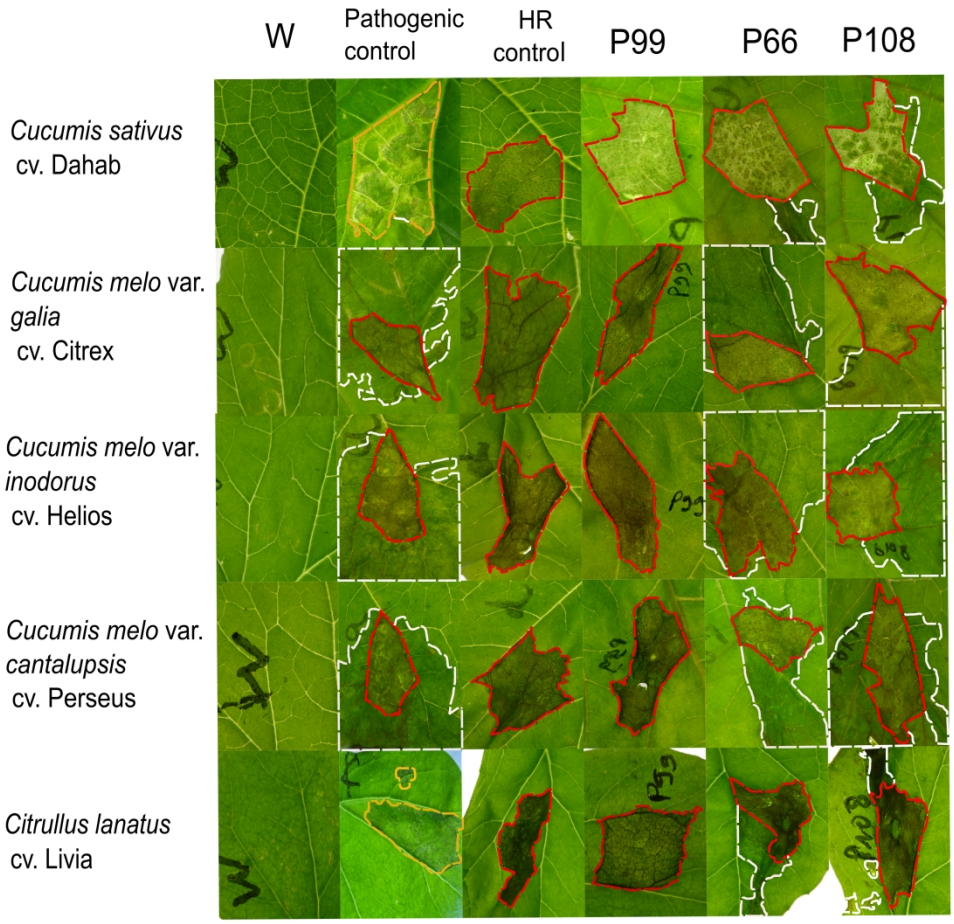


Fig. 4. Hypersensitive response of diverse cucurbits following infiltration of VCZ strains. Three VCZ strains (P99, P66 and P108) were infiltrated in leaves of four genotypes of Cucumis sp. and one Citrullus lanatus. Symptoms were pictured 72 h post-inoculation. Pathogenic controls consisted in A. citrulli CFBP 4459 on C. lanatus cv. Livia, P. syringae pv. aptata CFBP 5467 on melon cultivars and P. amygdali pv. lachrymans CFBP6463 on cucumber. The HR control for all tested plants was P. syringae pv. dysoxily CFBP 2356. Water was used as a negative control. Within the inoculated area the red line delimits necrosis and the orange line delimits water-soaked lesion. The white line delimits the expansion of water-soaked symptoms beyond the inoculation point.

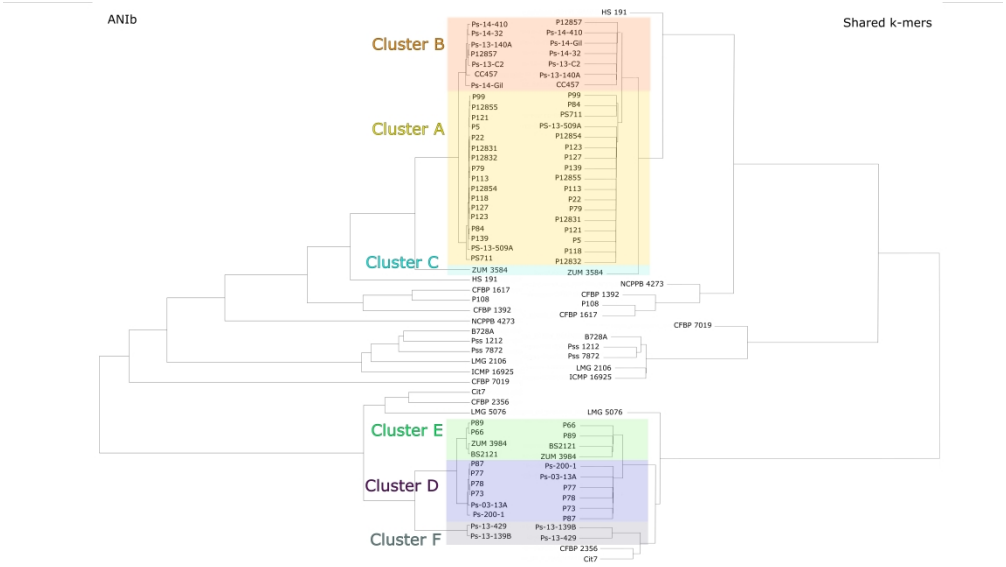


Fig. 5. Relationships between genomic sequences of P. syringae from phylogroup 2 using ANIb and k-mers approaches. Dendrograms were built on distances calculated from ANIb and k-mers matrices obtained with Ki-S (Briand et al. 2019). The color blocks represent the cucurbit clusters highlighted with both approaches.

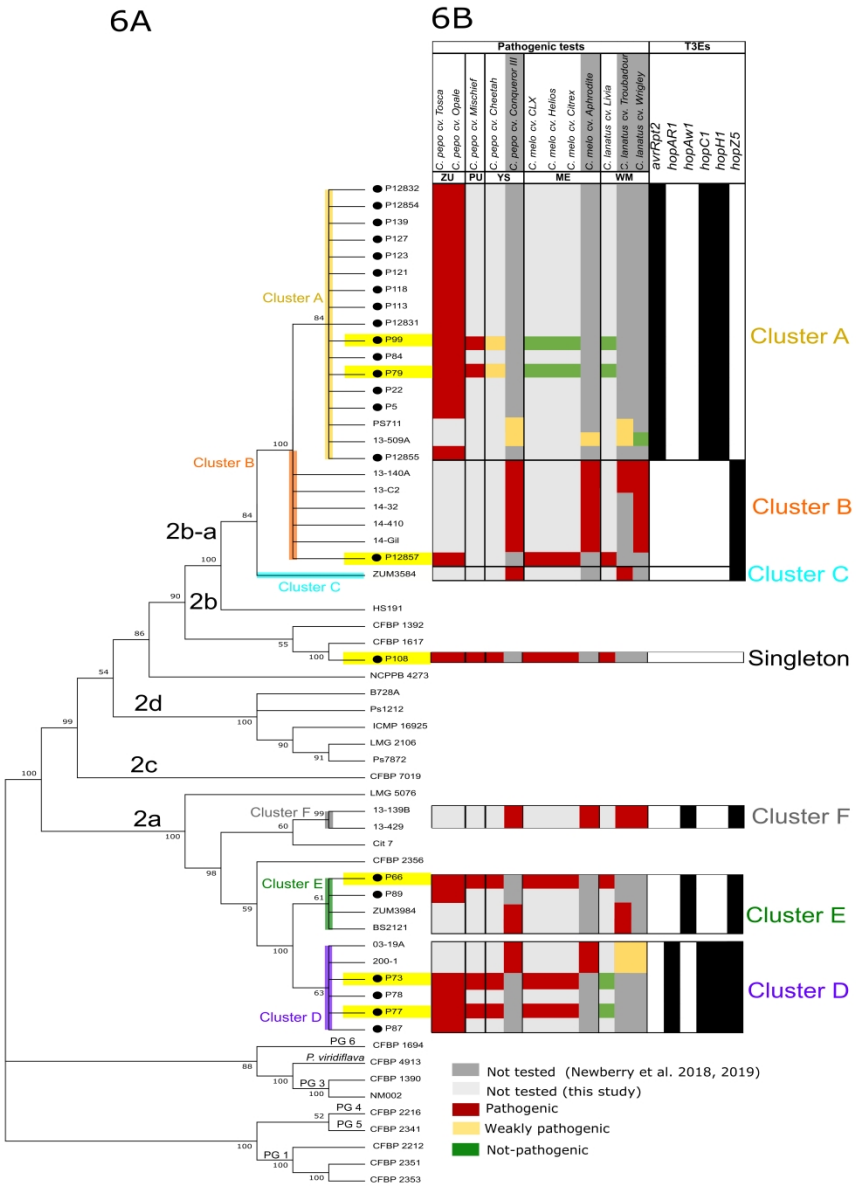


Fig. 6. Phylogeny of VCZ strains, host range and presence of some T3E encoding genes. (A) Phylogeny based on new MLSA scheme using seven genes (gapA, gltA, gyrB, rpoD, Psyr3420, Psyr4880 and Psyr3208). Maximum likelihood tree was based on concatenated partial sequences (4431 pb). Bootstrap scores (1,000 replicates) were displayed at each node. Black points indicate VCZ strains. (B) Summarized results obtained from pathogenicity tests and presence of differential genes encoding T3Es recovered from genome sequence analysis. Columns in dark grey report results from Newberry et al. 2018, 2019 that were available for strains Ps711, 13-509A, 13-140A, 13-C2, 14-32, 14-410, 14-Gil, Zum3584, 13-139B, 13-429, Zum3984, Bs2121, 03-19A and 200-1. The light grey report the results obtained in this study. Colors are given as following: light or dark grey when a strain is not tested on a given host, red when a strain is pathogenic on a given host, yellow when a strain is weakly pathogenic on a given host and green when a strain is not pathogenic on a given host. A black box means a T3E presence. Cucurbit species and sub-species are reported as following, ZU for zucchini, YS for yellow Straightneck squash, PU for pumpkin, ME for melons and WM for watermelon.

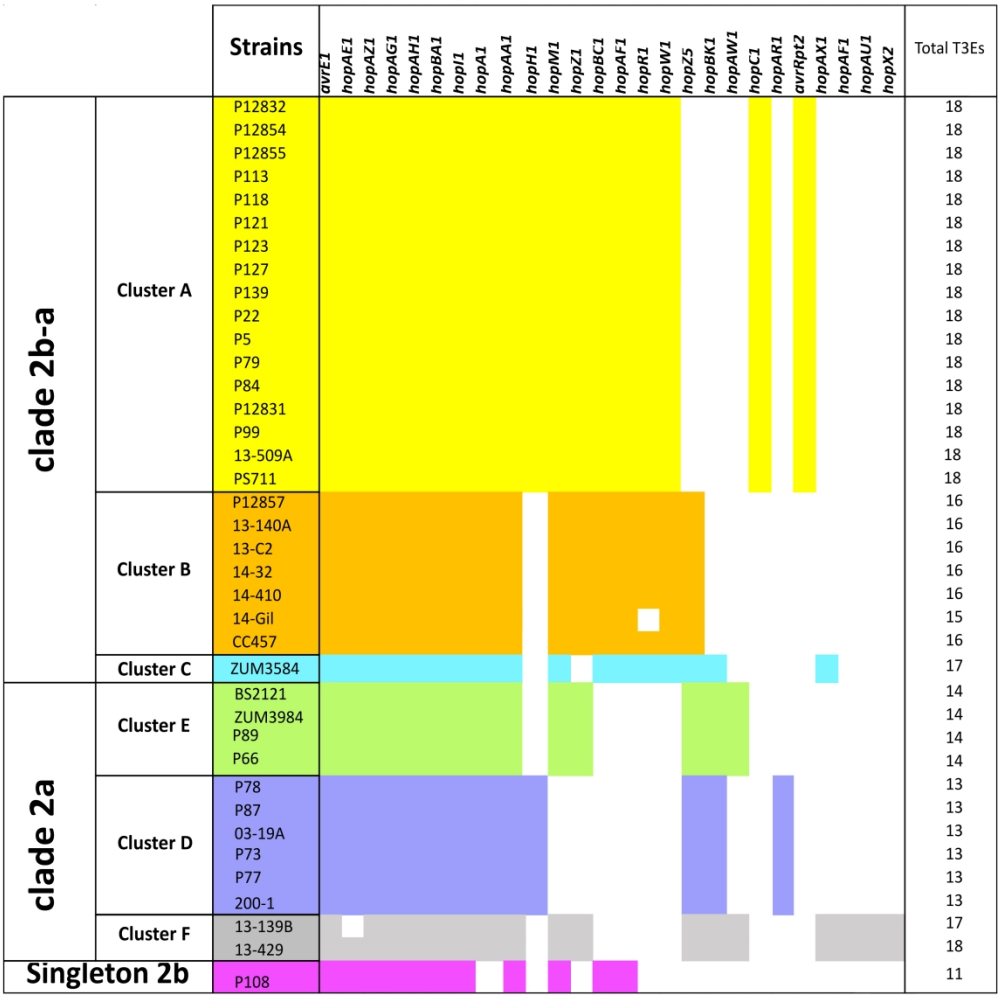
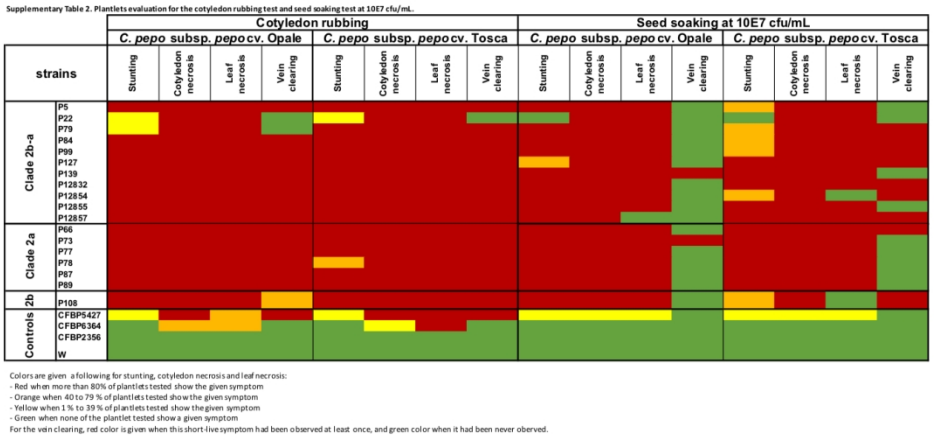
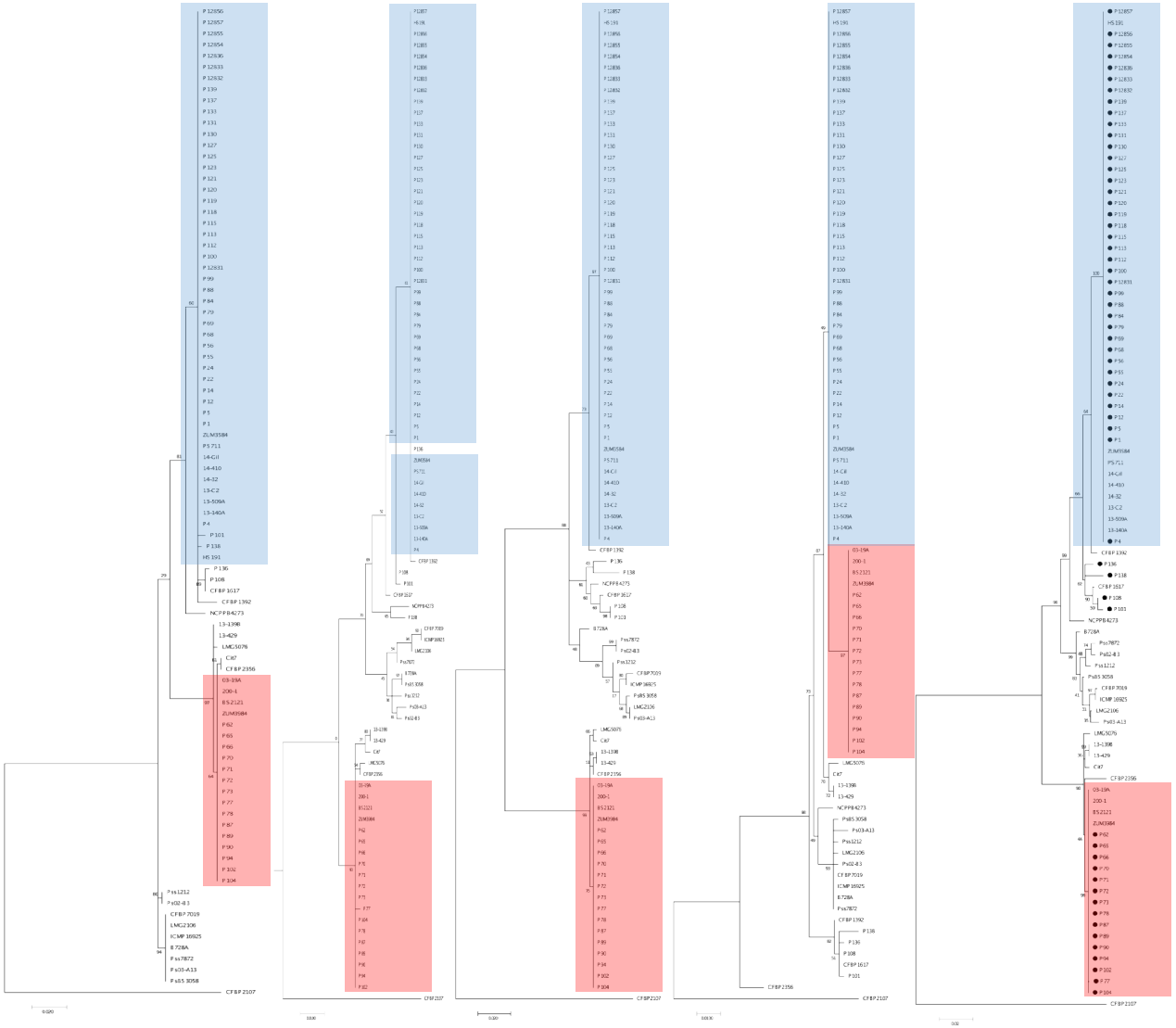


Fig. 7. Set of T3Es predicted in the genome sequences of VCZ strains and other strains isolated from cucurbits Color blocks reflect presence of the effector within a cluster (yellow: cluster A, orange: cluster B, blue: cluster C, green: cluster E, purple: cluster D, grey: cluster F, pink: singleton P108). White block means the T3E absence.



Supplementary table 2. Plantlets evaluation for the cotyledon rubbing test and seed soaking test at 1.107 cfu/mL. Colors refer to the incidence of a given symptom, with red for more than 80% of plantlets, orange 40 to 79%, yellow for 1 to 39% and green when none of the plantlets presented the symptom. For vein clearing, red color indicate that this symptom was observed at least at one date and green color when it had been never observed.

527x294mm (72 x 72 DPI)



gapA *gltA* *gyrB* *rpoD* Concatenated

Supplementary fig. 1
Phylogenies built on four individual housekeeping genes (*gapA*, *gltA*, *gyrB* and *rpoD*) and the concatenated dataset. Maximum-likelihood (ML) trees were generated with MEGA7 (Kumar et al. 2016) using the Tamura-Nei model with gamma distributed with invariant sites (G+I), and 1000 bootstrap replicates. Blue color represents the clade 2b-a and red color the cucurbit clade 2a. P values determined using the Shimodaira-Hasegawa test for congruence of tree topologies run on each tree based on four housekeeping genes, individually and the concatenated dataset. Values an asterisk are significantly depart from the hypothesis of congruency.

	gapA	*gltA*	*gyrB*	*rpoD*	Concat
gapA		0.006	0.000*	0.000*	0.002*
gltA	0.011*		0.000*	0.000*	0.028*
gyrB	0.011*	0.006*		0.000*	0.535
rpoD	0.000*	0.000*	0.000*		0.000*
Concat	0.027*	0.027*	0.269	0.003*	

P66-2A-E (disease)

P77-2a-D (HR)

P73-2a-D (HR)

P79-2ba-A (HR)

CFBP 2365 (HR control) (*P. syringae* pv. *dysoxyli*)

water

CFBP 4459 (pathogenic control) (*A. citrulli*)

Supplementary Fig. 2 | Response of watermelon (*C. lanatus* cv. Livia) to infiltration of *P. syringae* VCZ strains P66, P73, P77 and P79

Cucumis melo var. *cantalupsis* cv. Perseus

Citrullus lanatus cv. Livia

Cucumis sativus cv. Dahab

Cucumis melo var. *inodorus* cv. Helios

Cucumis melo var. *galia* cv. Citrex

Supplementary fig. 4. Response of melons, cucumber and watermelon to infiltration of strain P12857, responsible for VCZ from clade 2b -a .