



Genetic diversity and population structure of *Pseudomonas savastanoi*, an endemic pathogen of the Mediterranean area, revealed up to strain level by the MLVA assay

Yaseen Jundi Rahi^{1,2} · Silvia Turco¹ · Maria Claudia Taratufolo¹ · Marta Tati¹ · Matteo Cerboneschi³ · Stefania Tegli³ · Franco Valentini² · Annamaria D'Onghia² · Nicola Sante Iacobellis⁴ · Giorgio Mariano Balestra¹ · Angelo Mazzaglia¹

Received: 12 December 2019 / Accepted: 3 June 2020
© Società Italiana di Patologia Vegetale (S.I.Pa.V.) 2020

Abstract

Pseudomonas savastanoi is a bacterial species included in the *Pseudomonas syringae* complex. It is further sub-typed in pathovars which are the causal agents of a group of diseases of woody plants, such as the “knot disease” on olive and oleander and the bacterial canker on ash. Given its long-established presence in the Mediterranean area, the pathogen causing the aforementioned diseases can be considered endemic. Here, an MLVA approach was developed to assess the genetic relationships among and within those pathovars, with a specific focus on *P. savastanoi* pv. *savastanoi*. By analyzing the genome of the *P. savastanoi* pv. *savastanoi* strain NCPPB 3335 (accession n° CP008742), 14 Tandem Repeat (TR) loci were identified and the corresponding primers were designed and used for the amplification of genomic DNAs from 84 strains belonging to *Pseudomonas savastanoi* pathovars. Data were analyzed using different approaches, such as hierarchical clustering, STRUCTURE, and k-means clustering with DAPC to evaluate the effectiveness of the assay in describing pathovars and population structure of the pathogen. Results reveal a very complex articulation of genetic relationships, as expected from a long-time evolving pathogen, while providing the possibility to discriminate the pathovars between each other. At intra-pathovar level, the MLVA assay clusters isolates mainly according to their hosts and geographic origin. This resulted particularly useful in the identification and tracking of *P. savastanoi* populations at local level.

Keywords Knot disease · VNTR · MLVA · DAPC · Population structure

Introduction

The term “knot disease” in plant pathology describes the excessive growth of plant tissues or organs that leads to

hyperplastic and/or hypertrophic masses in the shape of tumors or knots. These formations can be clearly observed on the aerial parts of diseased plants mostly on the woody parts of the plant such as stem, branches, or twigs and, more rarely, on leaves and fruits. Generally, the disease is caused by the Gram-negative bacterium *Pseudomonas savastanoi* (Janse 1982; Iacobellis 2001; Ramos et al. 2012; Buonauro et al. 2015), which is a monomorphic pathogenic gamma-proteobacterium species belonging to the genomospecies 2 of the *Pseudomonas syringae* complex (Gardan et al. 1992). To date, six pathovars have been described within this species. *Pseudomonas* pv. *savastanoi* (Gardan et al. 1992), pv. *fraxini*, pv. *nerii* (Janse 1982), and pv. *retacarpa* (Garcia De Los Rios 1999) attack members of different plant families including Oleaceae (*Olea europaea*, *Jasminum officinalis*, *Forsythia* spp., *Phyllirea* spp., and *Ligustrum japonicum*), Fabaceae (*Retama*), Rhamnaceae (*Rhamnus* spp.), Myrtaceae (*Myrtus communis*), Apocinaceae (*Nerium oleander*, *Mandevilla sanderi*), and Lythraceae

✉ Angelo Mazzaglia
angmazza@unitus.it

¹ Department of Agricultural and Forestry Sciences, University of Tuscany, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy
² CIHEAM-Mediterranean Agronomic Institute of Bari, Via Ceglie 9, 70010 Valenzano, Ba, Italy
³ Agrifood Production and Environmental Sciences (DISPAA), University of Florence, Piazzale delle Cascine 18, 50144 Florence, Italy
⁴ School of Agricultural, Forestry, Food and Environmental Sciences, University of Basilicata, Viale dell'Ateneo Lucano, 85100 Potenza, Italy

(*Punica granatum*) (Bradbury 1986; Iacobellis et al. 1998; Marchi et al. 2011; Moretti et al. 2017). The pathovar *phaseolicola* (Gardan et al. 1992) is responsible for the halo blight of beans (*Phaseolus vulgaris* L.) (Völksch and Weingart 1997), with the exception of some strains of pv. *phaseolicola* that attack mulberry (Dworkin 2006), while pv. *glycinea* (Young 1978) is the causal agent of bacterial halo blight of soybean (*Glycine max*) and is restricted to this host species. *Pseudomonas savastanoi* pathovars are classified based on host range and genetic information (Mugnai et al. 1994). Early studies have focused on the differentiation and clarification of different pathovars using an array of morphological, serological, physiological, nutritional, biochemical, numerical taxonomy and DNA-DNA hybridization (Janse 1982; Gardan et al. 1992). These methods have shown some limitations over time, especially when differentiation or relatedness are required at strain level (Bennasar et al. 2010). The movement of goods around the world and the open borders have facilitated the diffusion of plant diseases through the movement of pathogens into new areas. Furthermore, climate change has helped the emergence or re-emergence of new pathogens, which negatively affect the food production (Grünwald and Goss 2011). Therefore, population genetic analyses are necessary to understand how a pathogen emerges and adapts to the environment (Grünwald et al. 2017). Indeed, the comprehension of population genetics has been enhanced by coalescent theory, computational methods and molecular biology, which have initiated the genomic era (Hartl and Clark 2007).

Only a few tools have been developed to investigate in detail and understand the genetic diversity among the strains of *Pseudomonas savastanoi* pathovars. Therefore, there is a necessity for a reliable method that can accurately discriminate between bacterial strains for crop surveillance, outbreak investigation and for the study of the evolution of plant pathogens in order to establish disease control strategies (Cesbron et al. 2014). Molecular typing methods are fast and powerful tools that allow to distinguish between closely related strains in an epidemiological survey, to determine relationships among the strains and to track their origin and pathways of spread. Molecular approaches can also help understanding evolution, host adaptation, and genetic diversity (Van Belkum et al. 2001; Scortichini et al. 2004; Wolska and Szwe 2012). Furthermore, they could reveal biological features of the pathogen that can affect virulence, pathogenicity, and host specificity, helping us to set up sustainable control measures (Tibayrenc 2005; Vinatzer et al. 2014).

The Restriction fragment Length Polymorphism (RFLP) has been one of the first molecular approaches to investigate *Pseudomonas savastanoi*; it has been used on a group of Italian strains from olive, oleander and ash trees for the evaluation of the differences among these strains (Sisto et al. 2002). PCR and random amplification of polymorphic DNA (RAPD) have also been used to evaluate associations between

strains of *Pseudomonas savastanoi* pathovars and their geographical distribution (Scortichini et al. 2004; Krid et al. 2009). Furthermore, fluorescent Amplified Fragment Length Polymorphism (f-AFLP) analysis has been used to appreciate genetic variability among a pool of representative strains from different olive regions around the world (Sisto et al. 2007). In Spain, the genetic diversity of *Pseudomonas savastanoi* pv. *savastanoi* has been investigated using the Insertion element (IS53) as typing method (López et al. 2008).

Despite the good results obtained, these methods remain time-consuming and poorly reproducible, limiting severely the possibility to compare results across laboratories (Cangelosi et al. 2004; Zhao et al. 2012). MLST is one of the most acknowledged methods for subtyping bacterial plant pathogens, though Sanger sequencing is always required. The results obtained by MLST were reported in agreement with those by DNA-DNA hybridization in *P. syringae* complex (Cinelli et al. 2014). More recently, Moretti and colleagues studied the genotypic diversity of Mediterranean populations using repetitive element palindromic PCR (rep-PCR) and MultiLocus Sequence Typing (MLST) of four housekeeping genes, namely *gap*, *gltA*, *gyrB*, and *rpoD* (Moretti et al. 2017). However, the sequences of the housekeeping genes usually show scarce diversity between individuals at subspecies level, which could result in failure when the aim is to resolve the evolutionary patterns of the bacterial populations (Achtman 2008), or to understand the populations pattern of highly homogeneous bacterial pathogens (Grissa et al. 2008).

In recent years Whole Genome Sequencing (WGS) and Single Nucleotide Polymorphism (SNPs) have been widely applied on many plant pathogenic bacteria, including *Pseudomonas savastanoi* pathovars (Pérez-Martínez et al. 2008; Rodríguez-Palenzuela et al. 2010; Moretti et al. 2014; Bartoli et al. 2015). These methods provide details about the entire genome sequences, but their cost is high; they also require advanced next-generation sequencing technologies, and most of all, high knowledge and computational skills, as well as devoted infrastructural facilities (Poulin et al. 2015; Davis et al. 2016).

MLVA (Multiple Loci Variable number of tandem repeat Analysis) is a very efficient typing technique. This molecular method is based upon the calculation of Variable copy Numbers of Tandem Repeats (VNTR). These tandemly repeated sequences (TR loci) have been found scattered throughout the prokaryotic and eukaryotic genomes (van Belkum et al. 1998; Lindstedt 2005); they are among the most variable regions in bacterial genomes, thus having the potential to resolve the genetic diversity of monomorphic pathogens. TRs occur in the genome due to a mismatch in DNA strands during replication (slipped-strand mispairing) or DNA recombination error (Gemayel et al. 2010; Vieira et al. 2016). TRs have been found to play an important role in bacterial adaptation (Saunders et al. 2000), phase variation (Gemayel

et al. 2010), and speed of loss gene disorder (Rocha 2003); they also facilitate evolutionary changes (Oliveira et al. 2006; Reinart et al. 2018), as well as genome plasticity and variation (Treangen et al. 2009). In order to calculate these TRs, PCR primers are designed in the conserved regions around these TRs. Then, the differences in the dimension of the amplicons among the strains are measured by capillary electrophoresis and converted in integers by deducting the right and the left flanking regions from the PCR amplicons, and dividing the remaining length by the size of the corresponding repeat unit in that locus. The numeric data obtained can be exchanged among laboratories directly or via public MLVA database (Grissa et al. 2008).

The discriminatory power of MLVA is based on the differences in the number of repeats in each of several loci, as generated by amplification of genomic DNA, resulting in a high level of polymorphism (Li et al. 2009). Moreover, MLVA is a fast procedure for monitoring recent, local outbreaks of bacterial pathogens, providing insights about the relationships at microevolutionary level (Ee et al. 2017).

MLVA methodology was primarily developed for human (van Belkum et al. 1997; Noller et al. 2003; Marsh et al. 2006) and animal bacterial pathogens (Le Flèche et al. 2006). Successively, it was applied for the first time on the plant pathogenic bacteria *Xylella fastidiosa* (Della Coletta-Filho et al. 2001) and, after that, it has been widely used to investigate bacterial plant diseases, as in *Xanthomonas citri* pv. *citri* (Vernière et al. 2014; Pruvost et al. 2014), *Xanthomonas arboricola* pv. *pruni* (Bergsma-Vlami et al. 2012), *Pseudomonas syringae* pv. *maculicola* and *Pseudomonas syringae* pv. *tomato* (Gironde and Manceau 2012), *Xanthomonas oryzae* pv. *oryzicola* (Zhao et al. 2012), *Ralstonia solanacearum* (N'Guessan et al. 2013), *Clavibacter michiganensis* subsp. *michiganensis* (Zaluga et al. 2013), *Erwinia amylovora* (Bühlmann et al. 2014), *Pseudomonas syringae* pv. *actinidiae* (Ciarroni et al. 2015), and *Ralstonia solanacearum* (Guinard et al. 2017).

In this study, we identified for the first time VNTRs loci and designed an MLVA assay on *Pseudomonas savastanoi*, to be analyzed by hierarchical clustering, k-means clustering and STRUCTURE software. The aim was to recognize pathovars and understand genetic relationships between populations within a wide collection of strains including worldwide isolates, isolates from an olive orchard from one region and a single tree of the same orchard.

Materials and methods

Bacterial strains and growth conditions

The study was carried out on a group of 84 strains of *Pseudomonas savastanoi* representative of the pathovars

savastanoi, *neri*, *fraxini*, *retacarpa*, and *phaseolicola*. Sixty-two strains were obtained from the DAFNE collection of the University of Tuscia in Viterbo, (Italy) and from the CIHEAM collection in Bari (Italy). Twenty-four additional samples were isolated from an olive orchard in Viterbo; 14 of them were collected randomly in the orchard, while the remaining 10 were obtained from a single tree. The bacteria were grown on King's B medium (KB) at 28 °C for 48 h before DNA extraction.

Genomic DNA extraction

Genomic DNA was extracted from freshly grown colonies using the QIAGEN kit (GmbH, Germany) following the manufacturer instruction. DNA quantification was obtained by a Qubit Fluorometer (Invitrogen, Life Technologies Italia, Monza, Italy), then its concentration was adjusted to 40 ng/μl with TE (10mMTris-HCl, 1 mM EDTA) buffer at pH 8.0, before storage at -20 °C until use. The identity of all the strains as *Pseudomonas savastanoi* was confirmed through *iaaL* gene amplification (Penyalver et al. 2000).

Tandem repeats identification and design of VNTR primers

The complete genome sequence of the *Pseudomonas savastanoi* pv. *savastanoi* strain NCPPB 3335 (GenBank accession number CP008742) was analyzed to ascertain the presence of candidate VNTR loci using Tandem Repeats Finder program (TRF; <https://tandem.bu.edu/trf/trf.html>) (Benson 1999) with the following parameters: tandem repeat range from 50 to 1000 bp, repeat unit length from 5 to 300 bp and identity higher than 80% between the copies of the tandem repeat array.

The VNTRs matching the predicted polymorphism size among *savastanoi*, *neri*, and *fraxini* pathovars were selected. To confirm the presence of these regions in all the available WGS sequences, both the right and left flanking regions of about 100 bp each were tested by BLAST (Altschul et al. 1990). Primer pairs were designed by Primer3plus software (<http://www.bioinformatics.nl/cgi-qbin/primer3plus/primer3plus>) in the flanking regions of each selected tandem repeat.

PCR amplification, agarose gel electrophoresis

The selected VNTR loci were amplified using a C1000™ Thermal Cycler (Bio-Rad, USA). Each PCR reaction mix contained 12.5 μL 2X GoTaq® master mix (Promega, Madison, USA), 9.5 μL nuclease-free water and 1 μL (40 ng) of template DNA, and 1 μL of each primer corresponding to 10 μM concentration, in a final volume of 25 μL. All the PCR reactions were run with an initial

denaturation step of 5 min at 94 °C followed by 35 cycles at 94 °C for 30 s, 50 °C to 67 °C (depending on the primer pair) for 30 s at 72 °C for 2 min with a final extension step of 10 min at 72 °C. PCR products were separated on 1.5% agarose gels and visualized under UV upon Gel Red® Nucleic Acid to confirm positive amplification. A random selection of samples underwent Sanger sequencing to check the exactness of the number of tandem repeats calculation.

VNTR analysis by capillary electrophoresis

The amplicons were analyzed using a QIAxcel multi-capillary electrophoresis system (QIAGEN, Milan, Italy). A DNA High-Resolution gel cartridge and the OM800 method were used to estimate amplicons sizes below 600 bp, with the following run parameters: 10 s of sample injection time; 5 kV of sample injection voltage; 3 kV of separation voltage for 700 s of separation time. The OM500 method was instead used for amplicons larger than 600 bp, with higher separation voltage (5 kV) and shorter separation time (500 s). The final results were analysed and interpreted using the Screen gel software (QIAGEN), which estimates both size and concentration of the amplicons.

Data analysis

The data matrix containing the tandem repeats numbers for 15 different loci, among 84 isolates, was imported into R version 3.4.4 (R Core Team 2018) and transformed into a *genind* object using the R package *adegenet* 2.1.1 (Jombart 2008). All the analyses of population structure and genetic diversity were performed using R packages *poppr*, version 2.8.1 (Kamvar et al. 2014, 2015), *adegenet* 2.1.1 (Jombart 2008), and *ade4* version 1.7-13 (Dray and Dufour 2007). The *poppr* *clonecorrect*() function was used to discard duplicated genotypes and the biases they could induce, as well as to remove the clones while collapsing them into unique multilocus genotype (MLGs). To evaluate if the number of loci was enough to describe the diversity between individuals, the *genotype_curve*() function of the *poppr* package was used with 1000 random samplings.

A hierarchical clustering analysis was performed with *hclust*() function of the package *stats* (R Core Team 2018) to reveal the population structure, using Bruvo's distance and UPGMA algorithm. The result was visualized as a dendrogram with the R package *factoextra* version 1.0.5 (Kassambara and Mundt 2017). Furthermore, *poppr* *bruvo.boot*() function was used to calculate Bruvo's distance with bootstraps with a cut-off threshold of 80.

The population structure was also evaluated using STRUCTURE software version 2.3.4 with K number set from 2 to 10 and 10 iteration runs per each K (Pritchard et al. 2000). Additionally, Evanno method was applied to obtain the

optimal ΔK estimation (Evanno et al. 2005). STRUCTURE results were combined using CLUMPP algorithm through the *clumppExport*() function of the same package and visualized with the function *plotQ*().

Also, a Discriminant Analysis of Principal Components (DAPC) was carried out using the R package *adegenet* (Jombart et al. 2010). The Bayesian Information Criterion (BIC) and the function *xval.Dapc*() were used to evaluate the number of clusters to be selected and the correct number of principal components to retain.

Results

The search within the *Pseudomonas savastanoi* pv. *savastanoi* genome sequence NCPPB 3335 performed via TRF program resulted in about 110 candidate VNTR loci. The comparison of the candidate loci against the WGS sequences of other *P. savastanoi* strains available in NCBI showed that less than 30% of these loci varied in size. A total of 20 candidate VNTR loci having a match percentage higher than 90% were selected and respective primer pairs were consequently designed. The TRs characteristics, the sequence of the primers, the position of the amplicon along the reference genome of NCPPB 3335 and the putative gene functions, if relevant, are summarized in Table 1. The 20 VNTRs loci were amplified on the 84 isolates for a total of 1680 reactions and the amplicons were separated via the QIAxcel multi-capillary electrophoresis system. The entire procedure was repeated twice for each sample. Five out of 20 loci gave multiple products or lacked reproducibility, which prompted their exclusion from further analysis.

A total number of 129 alleles were detected in all the final 15 VNTR loci, ranging from only 2 alleles for TRsav8 and TRsav17 loci, to a maximum of 30 alleles in the locus TRsav16 (data not shown).

The final data matrix containing the number of repeats, imported on R and clone corrected, resulted in 84 individuals collapsing into 78 unique multilocus genotypes (MLGs), adequate to describe the diversity between individuals, as shown by the genotype accumulation curve in Fig. 1 (Kamvar et al. 2015). The *poppr*() function from the same R package was used to obtain the diversity indexes reported in Table 2. More specifically, besides the number of individuals (N) and MLG per population, the expected MLG (eMLG) represents the number of MLG at the lowest sample size. Simpson's Index (λ), which is a measure of the probability that two randomly selected genotypes are different from each other, indicates that the *Pseudomonas savastanoi* pv. *savastanoi* clade, the more populated in the data set, resulted to be also the most divergent. Genotype Evenness (E.5) measures the genotype abundances within a population. In this case, *Pseudomonas savastanoi* pv. *fraxini* and *Pseudomonas savastanoi* pv.

Table 1 Characteristics, putative gene function and primers used for the TRs analyzed in this study

Name of TRs	TR unit length (bp)	TRs length	start- end	Gene function	Primers
TRsav1	9	158	995,987 996,340	<i>Hypothetical protein</i>	FR:ATTTCTGAGCGTCTGTGT RV:ATTAAGTGTGATTCTTTC
TRsav2	6	152	1,659,360 1,659,565	<i>Tellurium resistance protein Tera</i>	FR: CTTGAACCGCTGGCAAAA RV:CCGAAACCGCGCTGGATTT
TRsav3	7	166	1,988,604 1,990,085	<i>Glycosyl transferase</i>	FR: ATCTGGTGGGTTTCATGACC RV: CTCTGCATAATCGTATCCCT
TRsav6	7	193	935,627 938,149	<i>Flavodoxin</i>	FR:CTGGTGGATAACCGTCAGGT RV:AGCTGATCGAGCAAGGACGT
TRsav7	8	179	2,469,543 2,469,753	<i>Transposase</i>	FR: CTTGCCCATCTTGTGCACTT RV:GGCAACGCGCAGGCTCTGGA
TRsav8	6	153	5,324,874 5,325,044	<i>DNA topoisomerase IV subunit A</i>	FR: CGACCGTGAACAGAACTG RV: ACCGCCAGATCGGTCACATA
TRsav10	8	169	2,528,494 2,528,694	<i>4-hydroxy-tetrahydrodipicolinate synthase</i>	FR: GATGCTGGCTGAGGGTTG RV: GCGCAGATGCCCTGTCGATC
TRsav11	6	175	2,659,925 2,660,123	<i>heme ABC transporter ATP-binding protein CcmA</i>	FR: GCTCAATCTGTTGTGGGTTG RV:CTGCTGCCCGGCCGACAAGG
TRsav13	24	130	1,219,712 1,219,889	<i>Hypothetical protein</i>	FR: AAGATTTGGTACGCCAGCAG RV: GGTTTTACAGGTGGCCTCAC
TRsav15	8	194	134,818 135,075	<i>unknown</i>	FR: TTTGATGATCAGCCTTCGTG RV: GGCGCGATGATGGAGCGG
TRsav16	6	133	259,976 260,396	<i>Hypothetical protein</i>	FR: GTCAGATGCTTTTGGCTTGA RV:TGGAGATCCCTTTATTAATG AC
TRsav17	6	187	495,079 495,294	<i>Phosphodiesterase</i>	FR: ACCTATGGCGTGGTTCGATAC RV: TTGTGTCGATCGTCATGATT
TRsav18	9	157	996,131 996,323	<i>Hypothetical protein</i>	FR: ATAACCAGTCCGCGAGCTAA RV: TGTTTCATGAGCTAGAGAAA
TRsav19	7	173	1,990,097 1,990,297	<i>GDP-mannose dehydrogenase</i>	FR: TAAGTCAGTTGCGAGCCTCA RV: GACTCCCGAAGGCA AGCGCG
TRsav20	8	149	3,346,599 3,346,771	<i>unknown</i>	FR: GGTTCAGTGCATCAAACCAG RV: ATGGGCGAGGGTTGCTGTTC

phaseolicola resulted to be the ones with the highest genotype diversity. Finally, the highest Nei's diversity, which is the average genetic diversity per locus, also defined as expected heterozygosity (H_{exp}), was measured for *Pseudomonas savastanoi* pv. *savastanoi*.

Hierarchical clustering

The hierarchical clustering performed using Bruvo's distance and UPGMA as the agglomerative algorithm, generated the dendrogram in Fig. 2, where only bootstrap values higher than 80% are shown. According to the intrinsic nature of the method, the number of clusters depends on the choice of the threshold distance, and thus, in this first step of the analysis, the

organization of the groups was observed as it is, keeping the bootstrap values as key points for the validation.

In the dendrogram, the two *Pseudomonas savastanoi* pv. *phaseolicola* strains emerge as very similar to each other, and contemporaneously the most distant from all the other strains with a 100% bootstrap support, thus representing the outgroup for the analysis. Regarding the other pathovars, all the Italian strains belonging to the pathovar *nerii* clustered together; the only exception being the strain ITM305, that was instead isolated in California, USA, back in 1981. The same occurs for the strains of pathovar *fraxini*, which again form a clear cluster without exclusion. A single strain (SUPP3085) isolated from olive in Japan and the only strain belonging to the pathovar *retacarpa* are loosely related to this cluster. Then, the

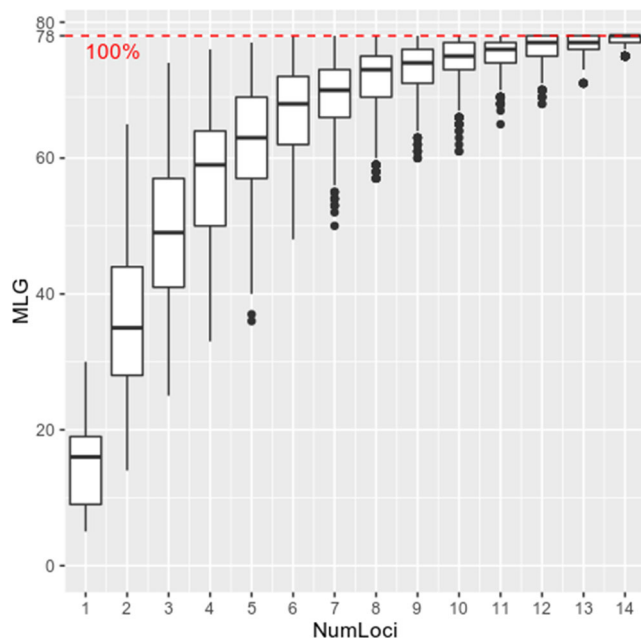


Fig. 1 Genotype accumulation curve for 84 isolates of *Pseudomonas savastanoi*. The x-axis represents the number of loci that was randomly sampled 1000 times up to (n-1) loci while the y-axis represents the number of multilocus genotypes observed up to 78 MLGs in the dataset. The dashed lines at the plateau indicate that enough number of loci were used to discriminate between individuals

numerous strains belonging to the pathovar *savastanoi* show a quite scattered distribution along the dendrogram, providing for a large genetic variability. Nonetheless, several clusters are recognizable, often in clear relationship with the plant host, when it is different from olive, or with their geographical origin. The cluster “*savast. C*”, formed by strains isolated from *Jasminum*, or the cluster “*savast. G*” whose strains were obtained from *Ligustrum*, are examples of liaison to host plant species, whilst geographic affinities referable to the country of isolation are evident for clusters “*savast. B*”-Albania, “*savast. E*”-Portugal, “*savast. F*”-Morocco, “*savast. H*”-Syria, and “*savast. C*”-California. An independent position in the dendrogram is coherently taken by single strains from Japan (SUPP3129 and SUPP3085) and from Tunisia (TN177). Regarding the Italian strains, it is worth noting the independent grouping of isolates from a single orchard in Viterbo

Table 2 *Poppr* function statistics: number of strains for each population (N), MultiLocus genotypes (MLG) and expected ones (eMLG), evenness (E), Simpon’s index (lambda) and Nei’s genetic diversity (expected heterozygosity)

<i>Pseudomonas savastanoi</i> pathovars	N	MLG	eMLG	E.5	lambda	Hexp
<i>savastanoi</i>	60	55	9.810206	0.905049	0.978889	0.640589
<i>fraxini</i>	11	11	10	1	0.909091	0.415931
<i>nerii</i>	10	9	9	0.9517	0.88	0.383704
<i>phaseolicola</i>	2	2	2	1	0.5	0.133333
<i>retacarpa</i>	1	1	1	NA	0	0
Total	84	78	9.88875	0.923356	0.985544	0.657298

Fig. 2 Dendrogram of 84 *Pseudomonas savastanoi* strains based on Bruvo’s distance and UPGMA as algorithm; the bootstrap values (for 1000 replicates) are given at the nodes

(Italy), among which those isolated from a single tree furtherly grouped together forming an additional subcluster.

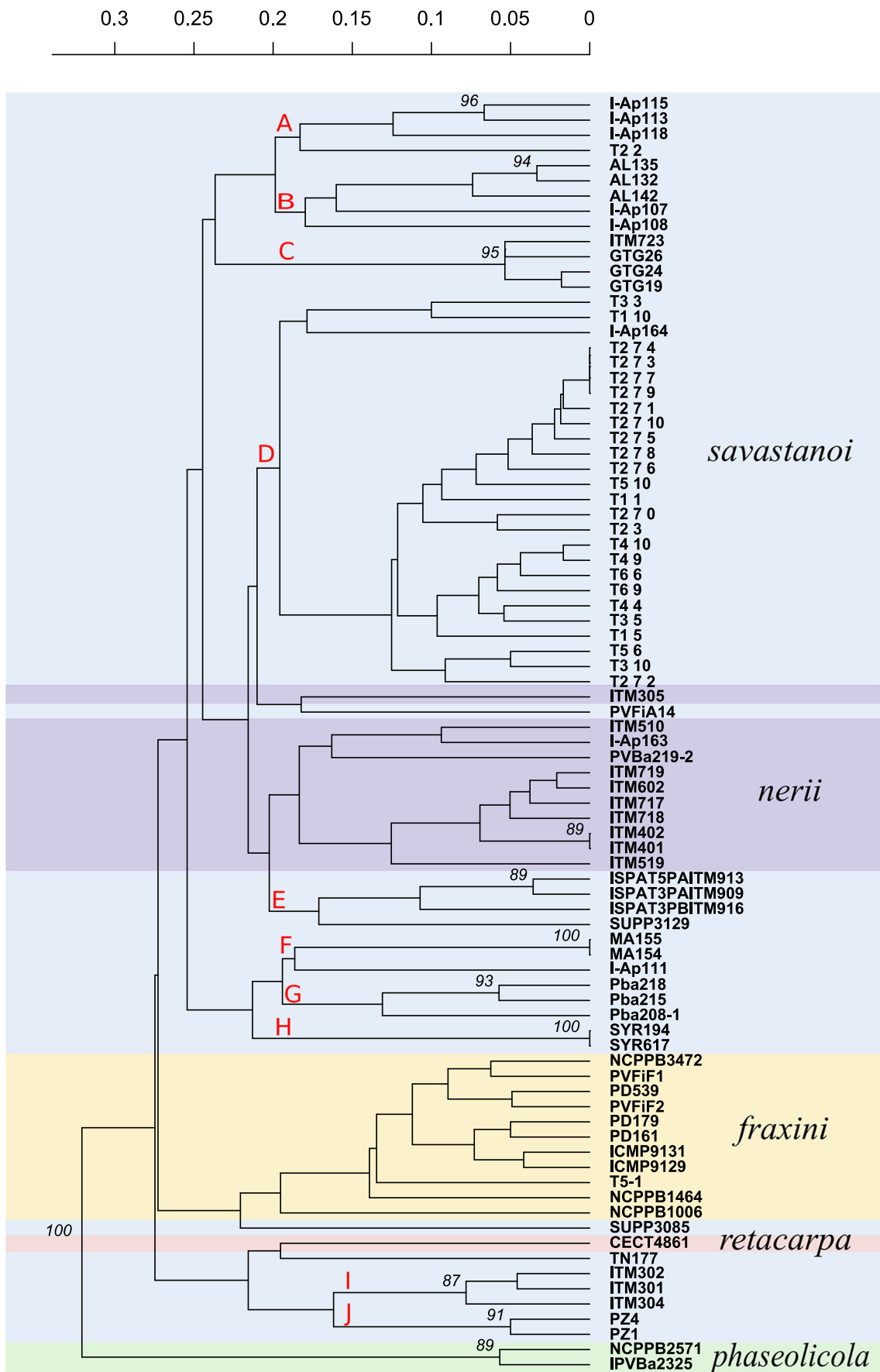
STRUCTURE analysis

The analysis of MLVA data using STRUCTURE software was run with K ranging from 2 to 12 and 10 independent iteration runs for each K. The results were imported in R environment and Evanno’s method was applied using *evannoMethodStructure()* function in the R package *pophelper* (Francis 2017) to obtain the optimal K value of 9, as showed in Fig. 3.

Thus, the 10 iterated results from K = 9 were combined and merged using CLUMPP and are shown in Fig. 4, where the percentage of assignment to each of the 9 groups is represented in the proportion of respective colors in individual bars. The previous hierarchical clustering is confirmed for the pathovar *fraxini* (Cluster 3 – dark pink), including the exception of strain NCPPB1006, and the pathovar *nerii* (Cluster 1 – orange), in which, however, two strains (ITM510 and PVBa219-2) show a deep admixture of possible clusters, as well as for the strain ITM305, already described out of this cluster in hierarchical clustering. It has to be noted that the two *phaseolicola* strains did not show a clear fitting to a single cluster as in the previous analysis. The STRUCTURE grouping instead reveals a strong signal of independence for the *savastanoi* strains from Viterbo (Cluster 7 – black), from Greece (Cluster 8 – gray-green), from Morocco and Syria (Cluster 9 – dark green), and from USA and Pomarico, Italy (Cluster 5 – light green). All the others show a less significant grouping admixture.

Discriminant analysis of principal components (DAPC)

The DAPC analysis in Fig. 5 also supported the hypothesis of grouping the strains in 9 clusters as the best solution, according to the Bayesian Information Criterion.



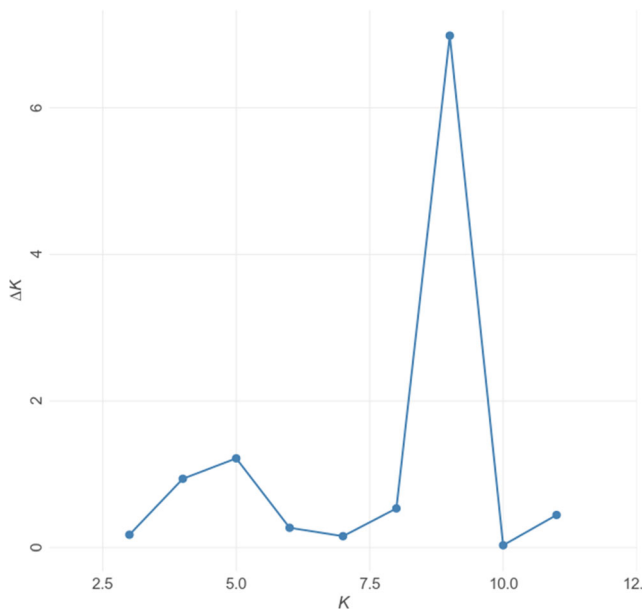


Fig. 3 Evanno’s method showing the number of K groups that best represent the population structure of *Pseudomonas savastanoi*

The resulting scatterplot in Fig. 6 shows the distribution of the clusters on LD axes when 15 principal components were retained, and second and third linear discriminants were selected. Once again, the general structure of groups as reported from previous analytical approaches is conserved. Indeed, the strains belonging to pathovar *fraxini* cluster separately from all the others (group 5 – yellow), except for the strain NCPPB 1006, along with pathovar *nerii* (group 3 – grey) and *phaseolicola* (group 2 – light blue). The numerous strains of the pathovar *savastanoi* are scattered in many groups, with a more or less admixed composition, which is represented by the overlapping ellipses, descriptive of the same subgroups identified before. The DAPC posterior probability of assignment for each individual to the 9 groups is shown in Figure 7.

Discussion

Molecular approaches are often used to define the evolutionary history, the host adaptation, and the genetic diversity of a pathogen (Van Belkum et al. 2001; Scortichini et al. 2004; Wolska and Szwe 2012). Furthermore, they could reveal biological features influencing virulence, pathogenicity, and host specificity of the pathogen, helping to define sustainable control measures (Tibayrenc 2005, 2009; Vinatzer et al. 2014). The mentioned molecular tools that allow the discrimination between individuals of the same taxon according to host preferences, as well as time and geographic origin, are of particular significance in understanding the spread dynamics of bacterial plant pathogens. In this context, an MLVA assay was tested for the first time to infer the population structure within

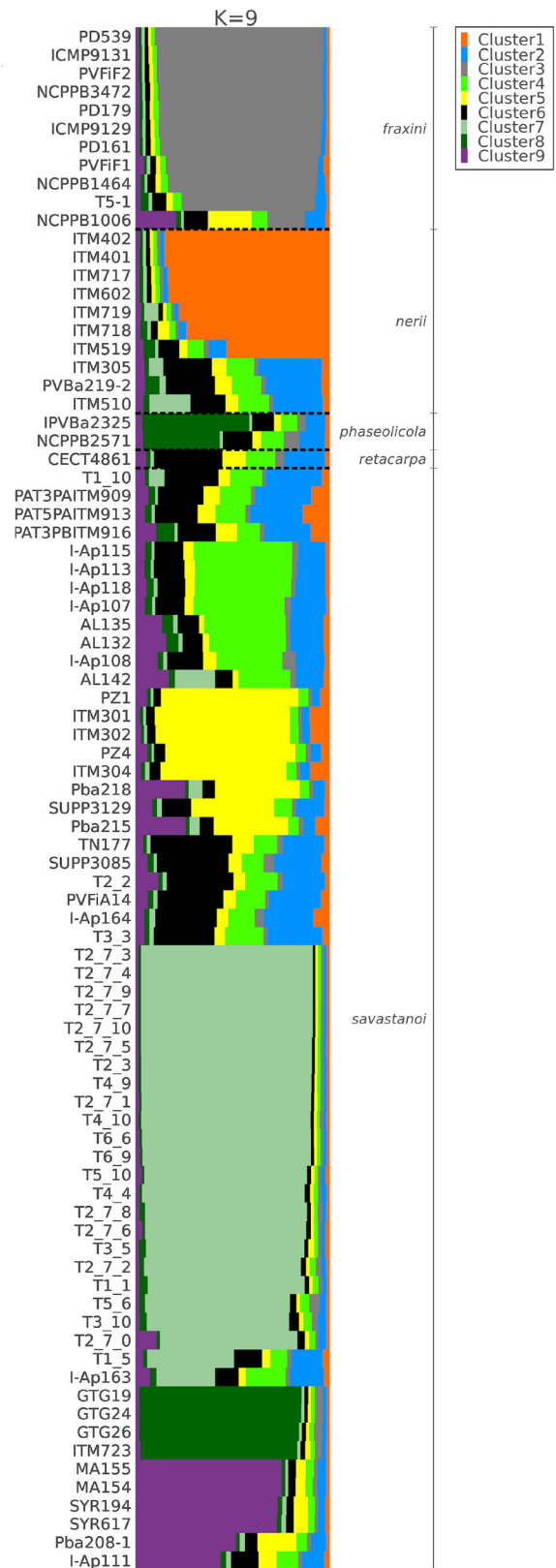


Fig. 4 Structure clump output for K = 9 following analysis of 84 isolates of *Pseudomonas savastanoi*, each strain represented by a single vertical line indicating its membership in each of K independent clusters. Pathovars indicated along the picture are separated by vertical black dashed lines

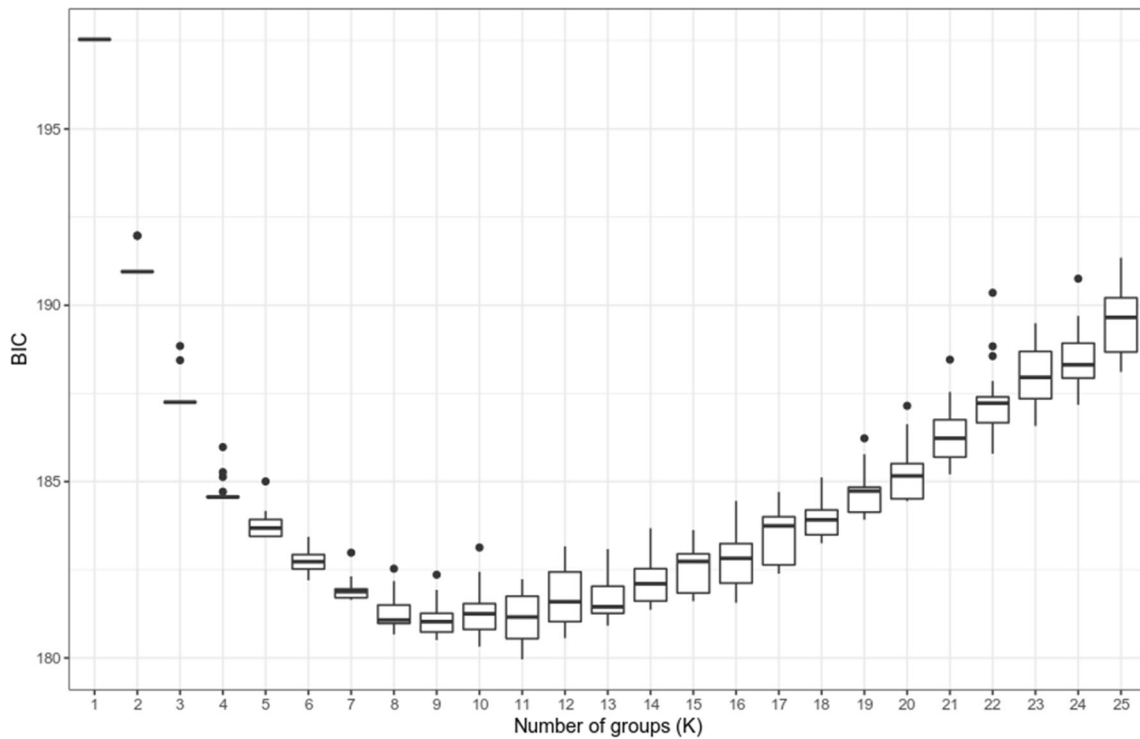


Fig. 5 Bayesian Information Criterion (BIC) curve for the k-means clustering of *Pseudomonas savastanoi* data set

Pseudomonas savastanoi pathovars. The 14 identified loci have proven to be able to provide enough information to disclose relationships among *Pseudomonas savastanoi* pathovars and strains.

As a first observation, the assay seems to be able to assign the samples to corresponding pathovars, as shown for the strains belonging to *phaseolicola*, *nerii*, and *fraxini* pathovars. Some exceptions, however, as the

scattered distribution of strains belonging to pathovar *savastanoi*, indicate that the high proficiency of MLVA in revealing very slight differences between individuals makes this method only partially suitable to gauge this type of information. Other molecular methods, such as MLST, are more appropriate for this type of phylogenetic information and already available to the scientific community (Pitt 2010).

Fig. 6 Scatterplot of genotypes in 9 groups according to DAPC. The individuals of the groups are distinguishable as in the present figure

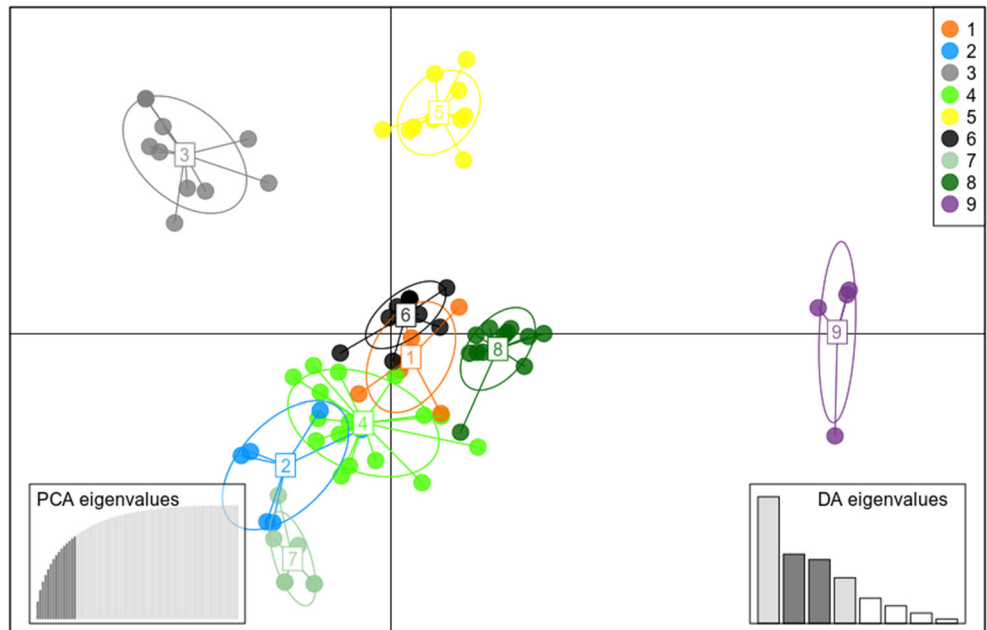
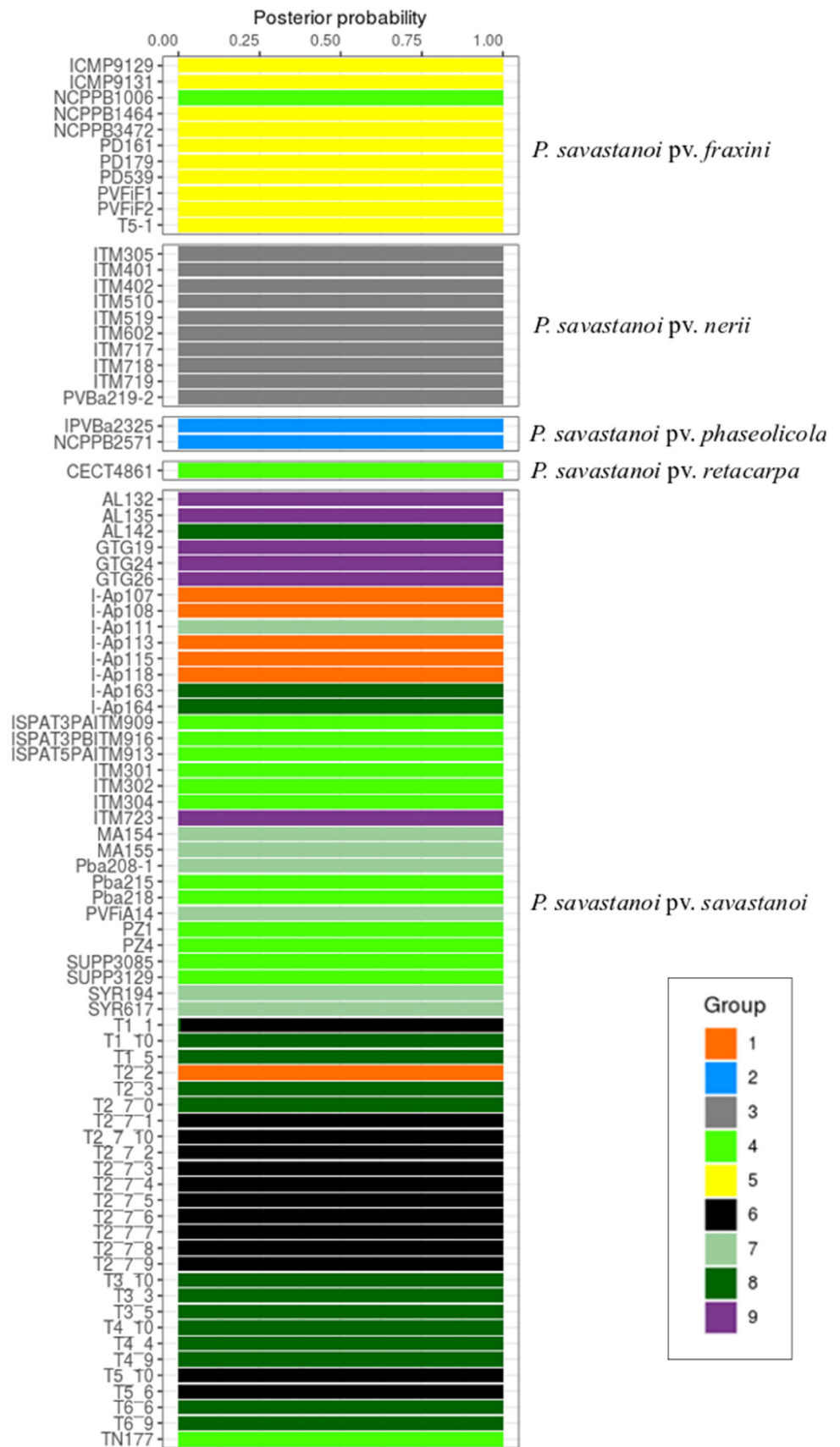


Fig. 7 The barplot shows the posterior probability of assignment of individuals in the 9 groups represented in different colours



When the data are analysed as a whole, the MLVA assay was instead able to provide information about population structure and to further separate strains in groups. Data elaboration was approached using three analytical methods and their congruence was investigated. Interestingly, the preliminary approach of STRUCTURE, through Evanno's method, and DAPC, through the BIC criterion, both recommended nine clusters as the best fitting to the variability of the input data set. This grouping was rather consistent between the two methods, in particular with the groups including strains of the pathovars *phaseolicola*, *fraxini* and *nerii* that were clearly defined, as well as with the group of strains isolated from a single orchard in Viterbo. About the other groups, STRUCTURE provided less clear results as demonstrated by strong admixture in the posterior assignment to these groups. DAPC, as expected from its ability to maximize variance between groups, assigned the individuals to the groups more sharply. The hierarchical clustering based on Bruvo's distance provided detailed results and a higher number of possible clusters. This suggests an interesting connection between clusters and both geographical origin of the strains and host plant species.

In any case, the overall framework depicted in this study indicates a large amount of genetic variability in *Pseudomonas savastanoi* in general and, more specifically, for the pathovar *savastanoi*. A first explanation resides in the heterogeneity of the collection analyzed, which included strains of worldwide origin, from 13 different countries, and isolated along a time span of more than 50 years. It is a further confirmation that MLVA, relying on genetic elements with a particularly high rate of mutation, finds its ultimate application in depicting more specific situations, such as, typically, clonal outbreaks and epidemics.

This leads to another aspect of this research, which concerns the analysis of strains isolated from knots concurrently collected from 10 olive plants in a small orchard (with a surface less than 1 ha wide) in Viterbo and from different knots of a single plant in the same orchard. In this scenario, if the disease would have been related to an outbreak from a single infective event, it would have been also logical to assume a substantial genetic homogeneity among its isolates. Conversely, in our case the level of genetic variability observed was still relatively high, even from knots collected on a single plant. This could in turn be explained by acknowledging that *P. savastanoi* pv. *savastanoi* is not an epidemic but an endemic pathogen, widespread in Italy for centuries. Such a long history of coevolution between the pathogen and its host plant has probably led to the differentiation of uncountable genotypes so that even on a single plant, as for plant T2.7, multiple infections by different genotypes commonly occurs. Furthermore, few isolates from other plants in the orchard are included in the cluster of the isolates from this single plant, just as few individuals from this single plant matched groups

of isolates from other plants. This great admixture on a very confined plot might be explained by a human-mediated transmission of the bacterial infection through infected tools during pruning from a plant to another.

In conclusion, this MLVA assay developed and applied for the first time to *Pseudomonas savastanoi*, has proven so responsive that even isolates from a single plant could be distinguished. Consequently, the most consistent goal of this method towards an endemic pathogen as *Pseudomonas savastanoi* pv. *savastanoi* should be the analysis of very specific situations. The method could be valuable to investigate biological facets such as how single genotypes infect and colonize tissues and organs of the same plant. Moreover, it could be used to evaluate possible interactions between strains colonizing the same tissue, or even to reconstruct micro-evolutive drifts over time in connection to changes in environmental conditions or in orchard management.

Acknowledgements The research was carried out in the frame of the MIUR (Ministry for education, University and Research) initiative "Department of Excellence" (Law 232/2016). We thank the CIHEAM Mediterranean Agronomic Institute of Bari for funding the PhD grant to Yaseen Jundi Rahi within HiSpro Project (High Studies Programme for Iraqi officials at Italian Academic Institutions) jointly with the Iraqi Ministry of Agriculture.

Compliance with ethical standards

Conflict of interest All the authors declare that they have no conflict of interest.

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu Rev Microbiol* 62:53–70. <https://doi.org/10.1146/annurev.micro.62.081307.162832>
- Altschul SF, Gish W, Miller W et al (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Bartoli C, Lamichhane JR, Berge O et al (2015) Mutability in *Pseudomonas viridiflava* as a programmed balance between antibiotic resistance and pathogenicity. *Mol Plant Pathol* 16:860–869. <https://doi.org/10.1111/mpp.12243>
- Bennasar A, Mulet M, Lalucat J, García-Valdés E (2010) PseudoMLSA: a database for multigenic sequence analysis of *Pseudomonas* species. *BMC Microbiol* 10:118. <https://doi.org/10.1186/1471-2180-10-118>
- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580. <https://doi.org/10.1093/nar/27.2.573>
- Bergsma-Vlami M, Martin W, Koenraadt H et al (2012) Molecular typing of Dutch isolates of *Xanthomonas arboricola* pv. *pruni* isolated from ornamental cherry laurel. *J Plant Pathol* 94:S1.29–S1.35. <https://doi.org/10.4454/jpp.v94i1sup.006>

- Bradbury JF (1986) Guide to plant pathogenic Bacteria; CAB international mycological institute. J Bradbury—1986—334 p 332
- Bühlmann A, Dreo T, Rezzonico F et al (2014) Phylogeography and population structure of the biologically invasive phytopathogen *Erwinia amylovora* inferred using minisatellites. *Environ Microbiol* 16:2112–2125. <https://doi.org/10.1111/1462-2920.12289>
- Buonaurio R, Moretti C, Da Silva DP et al (2015) The olive knot disease as a model to study the role of interspecies bacterial communities in plant disease. *Front Plant Sci* 6. <https://doi.org/10.3389/fpls.2015.00434>
- Cangelosi GA, Freeman RJ, Lewis KN et al (2004) Evaluation of a high-throughput repetitive-sequence-based PCR system for DNA fingerprinting of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex strains. *J Clin Microbiol* 42:2685–2693. <https://doi.org/10.1128/JCM.42.6.2685-2693.2004>
- Cesbron S, Pothier J, Gironde S et al (2014) Development of multilocus variable-number tandem repeat analysis (MLVA) for *Xanthomonas arboricola* pathovars. *J Microbiol Methods* 100:84–90. <https://doi.org/10.1016/j.mimet.2014.02.017>
- Ciaroni S, Gallipoli L, Taratufolo MC et al (2015) Development of a multiple loci variable number of tandem repeats analysis (MLVA) to unravel the intra-pathovar structure of *Pseudomonas syringae* pv. *actinidiae* populations worldwide. *PLoS One* 10:e0135310. <https://doi.org/10.1371/journal.pone.0135310>
- Cinelli T, Marchi G, Cimmino A et al (2014) Heterogeneity of *Pseudomonas savastanoi* populations infecting *Myrtus communis* in Sardinia (Italy). *Plant Pathol* 63:277–289. <https://doi.org/10.1111/ppa.12096>
- Davis EW, Weisberg AJ, Tabima JF et al (2016) Gall-ID: tools for genotyping gall-causing phytopathogenic bacteria. *PeerJ* 4:e2222. <https://doi.org/10.7717/peerj.2222>
- Della Coletta-Filho H, Takita MA, De Souza AA et al (2001) Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. *Appl Environ Microbiol* 67:4091–4095. <https://doi.org/10.1128/AEM.67.9.4091-4095.2001>
- Dray S, Dufour AB (2007) The *ade4* package: implementing the duality diagram for ecologists. *J stat Softw* 22:1–20. <https://doi.org/10.18637/jss.v022.i04>
- Dworkin M (2006) The prokaryotes - Vol. 6: Proteobacteria: gamma subclass. Springer, 3rd edn. Springer-Verlag New York, New York
- Ee CS, Ismail A, Banga Singh KK (2017) Chapter 3.7. Multiple-locus variable-number tandem repeat analysis (MLVA) for bacterial typing. In: sustainable diagnostics for low resource areas. Perpustakaan Negara, Malaysia
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Francis RM (2017) *pophelper*: an R package and web app to analyse and visualize population structure. *Mol Ecol Resour* 17:27–32. <https://doi.org/10.1111/1755-0998.12509>
- García De Los Ríos JE (1999) *Retama sphaerocarpa* (L.) Boiss., a new host of *Pseudomonas savastanoi*. *Phytopathol Mediterr* 38:54–60
- Gardan L, Bollet C, Ghorrah MA et al (1992) DNA relatedness among the pathovar strains of *Pseudomonas syringae* subsp. *savastanoi* Janse (1982) and proposal of *Pseudomonas savastanoi* sp. nov. *Int J Syst Bacteriol* 42:606–612. <https://doi.org/10.1099/00207713-42-4-606>
- Gemayel R, Vences MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet* 44:445–477. <https://doi.org/10.1146/annurev-genet-072610-155046>
- Gironde S, Manceau C (2012) Housekeeping gene sequencing and multilocus variable-number tandem-repeat analysis to identify subpopulations within *Pseudomonas syringae* pv. *maculicola* and *Pseudomonas syringae* pv. *tomato* that correlate with host specificity. *Appl Environ Microbiol* 78:3266–3279. <https://doi.org/10.1128/AEM.06655-11>
- Grissa I, Bouchon P, Pourcel C, Vergnaud G (2008) On-line resources for bacterial micro-evolution studies using MLVA or CRISPR typing. *Biochimie* 90:660–668. <https://doi.org/10.1016/j.biochi.2007.07.014>
- Grünwald NJ, Goss EM (2011) Evolution and population genetics of exotic and re-emerging pathogens: novel tools and approaches. *Annu Rev Phytopathol* 49:249–267. <https://doi.org/10.1146/annurev-phyto-072910-095246>
- Grünwald NJ, Everhart SE, Knaus BJ, Kamvar ZN (2017) Best practices for population genetic analyses. *Phytopathology* 107:1000–1010. <https://doi.org/10.1094/PHYTO-12-16-0425-RVW>
- Guinard J, Latreille A, Guérin F et al (2017) New multilocus variable-number tandem-repeat analysis (MLVA) scheme for fine-scale monitoring and microevolution-related study of *Ralstonia pseudosolanacearum* phylotype I populations. *Appl Environ Microbiol* 83. <https://doi.org/10.1128/AEM.03095-16>
- Hartl DL, Clark AG (2007) Principles of population genetics, fourth Edn. Sunderland, Massachusetts
- Iacobellis NS (2001) Olive knot. *Encycl Plant Pathol*:713–715
- Iacobellis NS, Caponero A, Evidente A (1998) Characterization of *Pseudomonas syringae* ssp. *savastanoi* strains isolated from ash. *Plant Pathol* 47:73–83. <https://doi.org/10.1046/j.1365-3059.1998.00202.x>
- Janse JD (1982) *Pseudomonas syringae* subsp. *savastanoi* (ex smith) subsp. nov., nom. Rev., the bacterium causing excrescences on *Oleaceae* and *Nerium oleander* L. *Int J Syst Bacteriol* 32:166–169. <https://doi.org/10.1099/00207713-32-2-166>
- Jombart T (2008) *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11:94. <https://doi.org/10.1186/1471-2156-11-94>
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2013:1–14. <https://doi.org/10.7717/peerj.281>
- Kamvar ZN, Brooks JC, Grünwald NJ (2015) Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front Genet* 6. <https://doi.org/10.3389/fgene.2015.00208>
- Kassambara A, Mundt F (2017) Package “factoextra” for R: extract and visualize the results of multivariate data analyses. R Packag. version, In <http://www.sthda.com/english/rpkgs/factoextra%0ABugReports>
- Krid S, Rhouma A, Quesada JM et al (2009) Delineation of *Pseudomonas savastanoi* pv. *savastanoi* strains isolated in Tunisia by random-amplified polymorphic DNA analysis. *J Appl Microbiol* 106:886–894. <https://doi.org/10.1111/j.1365-2672.2008.04058.x>
- Le Flèche P, Jacques I, Grayon M et al (2006) Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay. *BMC Microbiol* 6. <https://doi.org/10.1186/1471-2180-6-9>
- Li W, Raoult D, Fournier PE (2009) Bacterial strain typing in the genomic era. *FEMS Microbiol Rev* 33:892–916. <https://doi.org/10.1111/j.1574-6976.2009.00182.x>
- Lindstedt BA (2005) Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria.

- Electrophoresis 26:2567–2582. <https://doi.org/10.1002/elps.200500096>
- López MM, Quesada JM, Penyalver R et al (2008) Current technologies for *Pseudomonas* spp. and *Ralstonia solanacearum* detection and molecular typing. In: *Pseudomonas syringae* pathovars and related pathogens – identification. Epidemiology and Genomics, Springer Netherlands, pp 3–19
- Marchi G, Cinelli T, Surico G (2011) A review on *Pseudomonas savastanoi* genetic traits involved in disease development and in symptom induction. In: Schena L, Agosteo GE, Cacciola S.O. (eds) olive diseases and disorders. Chapter 6, pp 117–141. <https://doi.org/10.13140/2.1.2408.0328>
- Marsh JW, O’Leary MM, Shutt KA et al (2006) Multilocus variable-number tandem-repeat analysis for investigation of *Clostridium difficile* transmission in hospitals. J Clin Microbiol 44:2558–2566. <https://doi.org/10.1128/JCM.02364-05>
- Moretti C, Cortese C, da Silva DP et al (2014) Draft genome sequence of *Pseudomonas savastanoi* pv. *savastanoi* strain DAPP-PG 722, isolated in Italy from an olive plant affected by knot disease. Genome Announc 2. <https://doi.org/10.1128/genomeA.00864-14>
- Moretti C, Vinatzer BA, Onofri A et al (2017) Genetic and phenotypic diversity of Mediterranean populations of the olive knot pathogen, *Pseudomonas savastanoi* pv. *savastanoi*. Plant Pathol 66:595–605. <https://doi.org/10.1111/ppa.12614>
- Mugnai L, Giovannetti L, Ventura S, Surico G (1994) The grouping of strains of *Pseudomonas syringae* subsp. *savastanoi* by DNA restriction fingerprinting. J Phytopathol 142:209–218. <https://doi.org/10.1111/j.1439-0434.1994.tb04531.x>
- N’Guessan CA, Brisse S, Le Roux-Nio AC et al (2013) Development of variable number of tandem repeats typing schemes for *Ralstonia solanacearum*, the agent of bacterial wilt, banana Moko disease and potato brown rot. J Microbiol Methods 92:366–374. <https://doi.org/10.1016/j.mimet.2013.01.012>
- Noller AC, McEllistrem MC, Pacheco AGF et al (2003) Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157:H7 isolates. J Clin Microbiol 41: 5389–5397. <https://doi.org/10.1128/JCM.41.12.5389-5397.2003>
- Oliveira EJ, Pádua JG, Zucchi MI et al (2006) Origin, evolution and genome distribution of microsatellites. Genet Mol Biol 29:294–307
- Penyalver R, García A, Ferrer A et al (2000) Detection of *Pseudomonas savastanoi* pv. *savastanoi* in olive plants by enrichment and PCR. Appl Environ Microbiol 66:2673–2677. <https://doi.org/10.1128/aem.66.6.2673-2677.2000>
- Pérez-Martínez I, Zhao Y, Murillo J et al (2008) Global genomic analysis of *Pseudomonas savastanoi* pv. *savastanoi* plasmids. J Bacteriol 190:625–635. <https://doi.org/10.1128/JB.01067-07>
- Pitt TL (2010) Molecular epidemiology of microorganisms: methods and protocols. J Hosp Infect 75:247. <https://doi.org/10.1016/j.jhin.2010.02.002>
- Poulin L, Grygiel P, Magne M et al (2015) New multilocus variable-number tandem-repeat analysis tool for surveillance and local epidemiology of bacterial leaf blight and bacterial leaf streak of rice caused by *Xanthomonas oryzae*. Appl Environ Microbiol 81:688–698. <https://doi.org/10.1128/AEM.02768-14>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
- Pruvost O, Magne M, Boyer K et al (2014) A MLVA genotyping scheme for global surveillance of the citrus pathogen *Xanthomonas citri* pv. *citri* suggests a worldwide geographical expansion of a single genetic lineage PLoS One:9. <https://doi.org/10.1371/journal.pone.0098129>
- R Core Team (2018) R: A Language and Environment for Statistical Computing
- Ramos C, Matas IM, Bardaji L et al (2012) *Pseudomonas savastanoi* pv. *savastanoi*: some like it knot. Mol Plant Pathol 13:998–1009. <https://doi.org/10.1111/j.1364-3703.2012.00816.x>
- Reinar W, Knutsen J, Jentoft S et al (2018) Evolutionary adaptability linked to length variation in short genic tandem repeats. <https://doi.org/10.1101/310045>
- Rocha EPC (2003) DNA repeats lead to the accelerated loss of gene order in bacteria. Trends Genet 19:600–603. <https://doi.org/10.1016/j.tig.2003.09.011>
- Rodríguez-Palenzuela P, Matas IM, Murillo J et al (2010) Annotation and overview of the *Pseudomonas savastanoi* pv. *savastanoi* NCPPB 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. Environ Microbiol 12: 1604–1620. <https://doi.org/10.1111/j.1462-2920.2010.02207.x>
- Saunders NJ, Jeffries AC, Peden JF et al (2000) Repeat-associated phase variable genes in the complete genome sequence of *Neisseria meningitidis* strain MC58. Mol Microbiol 37:207–215. <https://doi.org/10.1046/j.1365-2958.2000.02000.x>
- Scortichini M, Rossi MP, Salerno M (2004) Relationship of genetic structure of *Pseudomonas savastanoi* pv. *savastanoi* populations from Italian olive trees and patterns of host genetic diversity. Plant Pathol 53:491–497. <https://doi.org/10.1046/j.0032-0862.2004.01051.x>
- Sisto A, Morea M, Baruzzi F, Palumbo G (2002) Differentiation of *Pseudomonas syringae* subsp. *savastanoi* strains isolated from various host plants by restriction fragment length polymorphism. Phytopathol Mediterr 41:63–71
- Sisto A, Cipriani MG, Tegli S et al (2007) Genetic characterization by fluorescent AFLP of *Pseudomonas savastanoi* pv. *savastanoi* strains isolated from different host species. Plant Pathol 56:366–372. <https://doi.org/10.1111/j.1365-3059.2007.01567.x>
- Tibayrenc M (2005) Bridging the gap between molecular epidemiologists and evolutionists. Trends Microbiol 13:575–580. <https://doi.org/10.1016/j.tim.2005.09.004>
- Tibayrenc M (2009) Microbial molecular epidemiology: an overview. Methods Mol Biol 551:1–12
- Treangen TJ, Abraham AL, Touchon M, Rocha EPC (2009) Genesis, effects and fates of repeats in prokaryotic genomes. In: FEMS Microbiology Reviews. pp 539–571. <https://doi.org/10.1111/j.1574-6976.2009.00169.x>
- van Belkum A, Scherer S, van Leeuwen W et al (1997) Variable number of tandem repeats in clinical strains of *Haemophilus influenzae*. Infect Immun 65:5017–5027
- van Belkum A, Scherer S, van Alphen L, Verbrugh H (1998) Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev 62:275–293
- van Belkum A, Struelens M, De Visser A et al (2001) Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. Clin Microbiol Rev 14:547–560. <https://doi.org/10.1128/CMR.14.3.547-560.2001>
- Vernière C, Bui Thi Ngoc L, Jarne P et al (2014) Highly polymorphic markers reveal the establishment of an invasive lineage of the citrus bacterial pathogen *Xanthomonas citri* pv. *citri* in its area of origin. Environ Microbiol 16:2226–2237. <https://doi.org/10.1111/1462-2920.12369>
- Vieira MLC, Santini L, Diniz AL, Munhoz C de F (2016) Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol 39:312–328. <https://doi.org/10.1590/1678-4685-GMB-2016-0027>
- Vinatzer BA, Monteil CL, Clarke CR (2014) Harnessing population genomics to understand how bacterial pathogens emerge, adapt to crop hosts, and disseminate. Annu Rev Phytopathol 52:19–43. <https://doi.org/10.1146/annurev-phyto-102313-045907>

- Völksch B, Weingart H (1997) Comparison of ethylene-producing *Pseudomonas syringae* strains isolated from kudzu (*Pueraria lobata*) with *Pseudomonas syringae* pv. *phaseolicola* and *Pseudomonas syringae* pv. *glycinea*. Eur J Plant Pathol 103:795–802. <https://doi.org/10.1023/A:1008686609138>
- Wolska K, Szwe P (2012) Genotyping techniques for determining the diversity of microorganisms. Genetic Diversity in Microorganisms. InTechOpen, In
- Young JM (1978) A proposed nomenclature and classification for plant pathogenic bacteria. New Zeal J Agric Res 21:153–177. <https://doi.org/10.1080/00288233.1978.10427397>
- Zaluga J, Stragier P, Van Vaerenbergh J et al (2013) Multilocus variable-number-tandem-repeats analysis (MLVA) distinguishes a clonal complex of *Clavibacter michiganensis* subsp. *michiganensis* strains isolated from recent outbreaks of bacterial wilt and canker in Belgium. BMC Microbiol 13. <https://doi.org/10.1186/1471-2180-13-126>
- Zhao S, Poulin L, Rodriguez-R LM et al (2012) Development of a variable number tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. Phytopathology 102: 948–956. <https://doi.org/10.1094/PHYTO-04-12-0078-R>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.