



## Article

# Molecular Epidemiology of *Xanthomonas euvesicatoria* Strains from the Balkan Peninsula Revealed by a New Multiple-Locus Variable-Number Tandem-Repeat Analysis Scheme

Taca Vancheva <sup>1,2</sup>, Nevena Bogatzevska <sup>3</sup>, Penka Moncheva <sup>2</sup>, Saša Mitrev <sup>4</sup>, Christian Vernière <sup>5</sup> and Ralf Koebnik <sup>1,5,\*</sup>

<sup>1</sup> IPME, Univ Montpellier, Cirad, IRD, Montpellier, France; tacavancheva@gmail.com

<sup>2</sup> Department of General and Industrial Microbiology, Faculty of Biology, Sofia University 'St. Kliment Ohridski', Sofia, Bulgaria; montcheva@biofac.uni-sofia.bg

<sup>3</sup> Institute of Soil Science, Agrotechnologies and Plant Protection 'Nikola Pushkarov', Sofia, Bulgaria; nbogatzevska@abv.bg

<sup>4</sup> Department for Plant and Environment Protection, Faculty of Agriculture, Goce Delchev University, Štip, North Macedonia; sasa.mitrev@ugd.edu.mk

<sup>5</sup> Plant Health Institute of Montpellier (PHIM), Univ Montpellier, Cirad, INRAe, Insitut Agro, IRD, Montpellier, France; christian.vernieres@cirad.fr

\* Correspondence: koebnik@gmx.de; Tel.: +33-467-416-228

**Citation:** Vancheva, T.; Bogatzevska, N.; Moncheva, P.; Mitrev, S.; Vernière, C.; Koebnik, R. Molecular Epidemiology of *Xanthomonas euvesicatoria* Strains from the Balkan Peninsula Revealed by a New Multiple-Locus Variable-Number Tandem-Repeat Analysis Scheme. *Microorganisms* **2021**, *9*, 536. <https://doi.org/10.3390/microorganisms9030536>

Academic Editor: Essaid Ait Barka

Received: 1 February 2021

Accepted: 1 March 2021

Published: 5 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Bacterial spot of pepper and tomato is caused by at least three species of *Xanthomonas*, among them two pathovars of *Xanthomonas euvesicatoria*, which are responsible for significant yield losses on all continents. In order to trace back the spread of bacterial spot pathogens within and among countries, we developed the first multilocus variable number of tandem repeat analyses (MLVA) scheme for pepper- and tomato-pathogenic strains of *X. euvesicatoria*. In this work, we assessed the repeat numbers by DNA sequencing of 16 tandem repeat loci and applied this new tool to analyse a representative set of 88 *X. euvesicatoria* pepper strains from Bulgaria and North Macedonia. The MLVA-16 scheme resulted in a Hunter–Gaston Discriminatory Index (HGDI) score of 0.944 and allowed to resolve 36 MLVA haplotypes (MTs), thus demonstrating its suitability for high-resolution molecular typing. Strains from the different regions of Bulgaria and North Macedonia were found to be widespread in genetically distant clonal complexes or singletons. Sequence types of the variable number of tandem repeats (VNTR) amplicons revealed cases of size homoplasy and suggested the coexistence of different populations and different introduction events. The large geographical distribution of MTs and the existence of epidemiologically closely related strains in different regions and countries suggest long dispersal of strains on pepper in this area.

**Keywords:** bacterial spot; genetic diversity; molecular typing; pepper; tomato

## 1. Introduction

Bacterial spot, which is caused by at least three *Xanthomonas* species, is a threatening disease of pepper and tomato plants worldwide [1–3]. Under favourable climatic conditions, the pathogens can cause significant yield losses in pepper- and tomato-growing areas. Bacteria enter via natural openings and colonize the apoplast. The disease is characterized by small, irregular, water-soaked, greasy-appearing lesions on all above-ground parts of the plants, followed by cell death, tissue necrosis and egress of *Xanthomonas* to the leaf surface [4]. Defoliation and shedding of fruits and blossoms are commonly observed in pepper production areas. This leads to a reduction in fruit quality and fruit loss due to the presence of lesions and the proliferation of secondary postharvest pathogens [5].

A wide range of genetic and physiological variation was found among *Xanthomonas* strains that cause bacterial spot on tomatoes and peppers. Until the early 1990s, the causal

agent of bacterial spot on pepper and tomato was *Xanthomonas campestris* pv. *vesicatoria*. In 1994, Stall et al. identified two genetically distinct groups of strains, referred to as A and B strains, within a worldwide collection based on genotypic and phenotypic characteristics [6]. Vauterin et al. proposed reclassification of the bacterial spot agents and divided *X. campestris* pv. *vesicatoria* into two species, with phenotypic group-A strains placed to *Xanthomonas axonopodis* pv. *vesicatoria* and phenotypic group-B strains relegated to *Xanthomonas vesicatoria*, as initially proposed by Dowson in 1939 [7,8]. In 2000, restriction fragment length polymorphisms and DNA sequence information of the 16S rRNA and adjacent intergenic sequence allowed to define two new pepper- and/or tomato-pathogenic *Xanthomonas* groups, C and D, with the D-group strains assigned to the species level *Xanthomonas gardneri* [9,10], which had been introduced in 1966 by Dye [11]. Based on DNA-DNA hybridization experiments, Jones et al. suggested in 2004 to consider the C-group strains as a novel species, called *Xanthomonas perforans* [12]. Only recently, genome-wide nucleotide sequence comparisons revealed that *X. perforans* does not deserve species status and was reclassified as a pathovar of *Xanthomonas euvesicatoria*, thus belonging to the same species as the previous group-A strains, also known as *X. axonopodis* pv. *vesicatoria*, which were reclassified as *X. euvesicatoria* pv. *euvesicatoria* [13–15]. Finally, based on whole-genome analyses, group-D strains, also known as *X. gardneri*, have been successively reclassified as *Xanthomonas cynarae* pv. *gardneri* and then as *Xanthomonas hortorum* pv. *gardneri* [16,17]. These new data corroborated earlier findings from partial sequencing of housekeeping genes that had suggested close phylogenetic relationships between *X. euvesicatoria* and *X. perforans* and between *X. gardneri* and *X. cynarae* [18–21].

Following this series of taxonomic reclassifications, bacterial spot of pepper and tomato is at present considered to be caused by three different species: *X. euvesicatoria*, *X. vesicatoria*, and *X. hortorum*. While *X. euvesicatoria* pv. *euvesicatoria* and *X. hortorum* pv. *gardneri* strains are well known to be pathogenic of both tomato and pepper plants, *X. vesicatoria* primarily infects tomato plants. *X. euvesicatoria* pv. *perforans* was considered to be restricted to tomato plants, until, in 2012, a strain was isolated from a pepper plant [22]. Historically, the two pathogens *X. euvesicatoria* pv. *euvesicatoria* and *X. vesicatoria* have had a worldwide distribution and were considered as the dominant bacterial spot lineages. More recently, however, *X. euvesicatoria* pv. *perforans* and *X. hortorum* pv. *gardneri* strains are increasingly often isolated in North and South America, Middle East, East Africa and regions bordering the Indian Ocean [23–32]. Perhaps this wide distribution in different geographical regions is due to introduction of contaminated seeds and/or seedlings [27,29,33,34]. Further, emergence of new strains or lineages through multiple recombination events and acquisition of novel transcription activation-like effector (TALE) may also contribute to this increasing distribution [35,36].

The European and Mediterranean Plant Protection Organization (EPPO) considers all bacterial spot pathogens of pepper and tomato as A2 quarantine pests, which means that they are locally present in the EPPO region [37,38]. For instance, the pathogens have been reported for Italy [39,40], for the Czech and Slovak Republics [41,42] and for the Balkan Peninsula [43]. In the East Balkan, bacterial spot of tomato and pepper was first reported in the 20th century, first on tomato in 1936 (Bulgaria) and then on pepper in 1965 for Bulgaria and in 1999 for North Macedonia [44–46]. Since then, the disease has become one of the economically most important diseases of pepper and tomato plants with losses reaching 10% to 20% [47,48]. *X. euvesicatoria* was identified as the dominant species of bacterial spot of pepper plants in Bulgaria and North Macedonia [49,50]. After 2014, the species was also reported as a bacterial spot agent of tomato in Bulgaria [51]. Analyses using restriction fragment length polymorphisms (RFLP), randomly amplified polymorphic DNA (RAPD) markers and rep-PCR revealed substantial genetic diversity among the Bulgarian and North Macedonian *X. euvesicatoria* strains [52,53]. Nevertheless, little information is at present available on the pathogen population structure in Bulgaria and North Macedonia.

Sustainable control measures, such as the use of chemicals, antagonists and/or resistant varieties, will critically depend on a profound knowledge of the population

structure and dynamics of the pathogen. For that aim, bacterial typing techniques are nowadays used for reliable and quick differentiation of closely related strains. Over the past decades, a variety of different typing methods have been developed to generate strain-specific patterns. However, only a limited number of molecular tools are able to distinguish strains at high resolution for epidemiological purposes. Whereas multilocus sequence analyses (MLSA) unambiguously identified bacterial spot agents at the species level, they did not emerge as a tool for differentiation of strains within a species due to very few sequence polymorphisms and signs of recombination events in some of the used housekeeping genes [21,27,29,54]. MLSA schemes targeting four to seven housekeeping genes may thus not allow drawing conclusions with respect to the global movement of the pathogens, and are certainly not useful for studying the population structure at a smaller geographical scale (within country) or characterize outbreak situations.

Multilocus variable numbers of tandem repeat analyses (MLVA) have become increasingly popular for high-resolution molecular typing of bacteria because of the high discriminatory power and reproducibility, ease of performance and portability, rapidity and low costs. This is correlated to the greater availability of bacterial genomes, which facilitated their development. Detection and analysis of polymorphic short sequence repeats organized at distinct loci showing high mutation rates has been proven to be a promising tool for epidemiological studies of monomorphic plant-pathogenic bacteria, such as *Pseudomonas syringae*, *Ralstonia solanacearum*, and several species of *Xanthomonas* [55–65].

Tandem repeat (TR) loci with small (<10 bp) repeat unit sizes are identified as microsatellites. Microsatellites alleles are usually defined from different sizes of DNA fragments revealed after electrophoresis. These electromorphs are then easily scored as repeat numbers, whose variation contributes to discriminate between individuals or populations. The numbers of repeats change through mutational events producing a variable number of tandem repeats (VNTR). These events are favoured by polymerase slippage at DNA replication corresponding to the addition or deletion of repeat motifs [66,67]. Thus, two alleles can share the same electromorph, i.e., are identical in state, following convergent mutational events, without originating from the same ancestral allele, i.e., are not identical by descent. This size homoplasmy, where alleles with identical sizes have different evolutionary histories, is related to the process of mutation. Two main models describe the process of mutations at TR loci. Under the stepwise mutation model (SMM), loss or gain of a single TR occurs with equal probabilities. In contrast, under the infinite-allele model (IAM), a new allelic state results from a unique mutational event involving any number of TRs, thereby not counting the actual number of gained or lost repeats [67]. A wealth of data suggested that most microsatellites evolve under an SMM [68]. Another type of variation can be generated within these microsatellite loci due to sequence variation occurring either in the tandem repeats or in the regions flanking the repeats. DNA sequence information can uncover cases of size homoplasmy that arise for instance from different combinations of different repeats within a compound microsatellite producing the same electromorph or from small mutations (InDels) within the flanking regions [67,69,70]. This detectable fraction corresponds to the molecularly accessible size homoplasmy (MASH).

In order to trace back the spread of bacterial spot pathogens within and among countries, we developed the first MLVA scheme for pepper- and tomato-pathogenic strains of *X. euvesicatoria* by directly assessing the repeat number by sequencing TR loci and applied this new tool to analyse a representative set of 88 *X. euvesicatoria* strains that were isolated from pepper plants in Bulgaria and North Macedonia.

## 2. Materials and Methods

### 2.1. Prediction of VNTR Loci and Primer Design

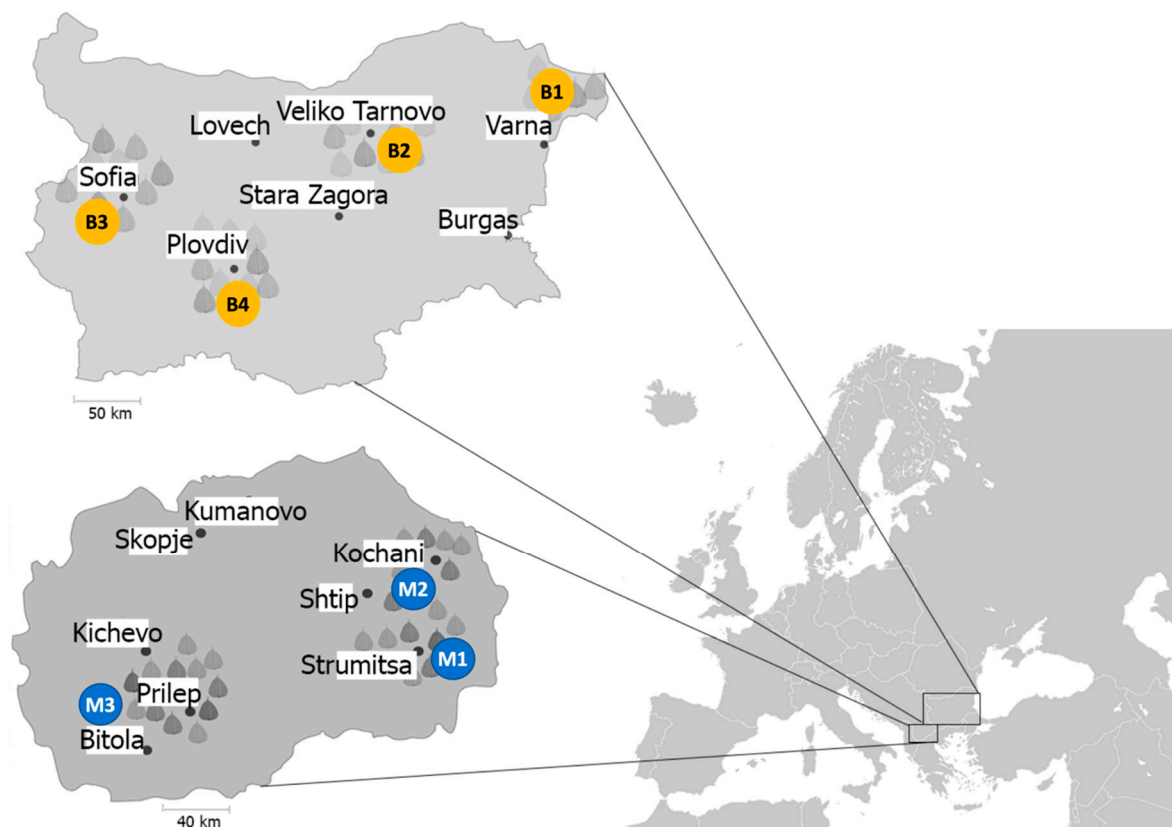
The complete genome sequence of *X. euvesicatoria* pv. *euvesicatoria* (also known as *Xanthomonas campestris* pv. *vesicatoria*) 85-10 (from pepper, GenBank accession number AM039952) and seven draft genome sequences of *X. euvesicatoria* pv. *euvesicatoria* strains,

including one North Macedonian strain (83M, from pepper, acc. no. JSZH00000000), one Bulgarian strain (66b, from pepper, acc. no. JSZG00000000), two Indian strains (LMG 905, from the unknown host plant, acc. no. JTEI00000000; LMG 918, from pepper, acc. no. JTEK00000000), one strain from Ivory Coast (LMG 909, from pepper, acc. no. JTEJ00000000), one strain from Brazil (LMG 933, from pepper, acc. no. JTEL00000000), and one strain from Tonga (LMG 667, from tomato, acc. no. JTEH00000000) [13,71,72] were screened for the presence of candidate VNTR loci using a bioinformatics pipeline, as previously described (<http://www.biopred.net/VNTR/>; accessed on 27 November 2020) [58]. In addition, draft genome sequences of representative strains of four other pathovars of *X. euvesicatoria*, all from the so-called Rademaker group 9.2 [73,74], were included: pv. *perforans* strain 91-118 (also known as *Xanthomonas perforans*; from tomato, acc. no. AEQW00000000), pv. *alfalfae* strain CFBP 3836 (also known as *Xanthomonas axonopodis* pv. *alfalfa*; from lucerne, acc. no. AUWN00000000), pv. *allii* strain CFBP 6369 (also known as *X. axonopodis* pv. *allii*; from onion, acc. no. JOJQ00000000), and pv. *citrumelo* strain F1 (also known as *X. axonopodis* pv. *citrumelo*; from citrus, acc. no. CP002914) [75–78].

For VNTR prediction, parameters were set as follows using the algorithm Tandem Repeats Finder (TRF) [79]: region length, 30 to 1000 bp; unit length, 5 to 9 bp; at least 6 copies and at least 80% similarity between adjacent repeats. Predicted VNTR loci were grouped by homology based on conservation of their 500-bp flanking sequences. Loci were named in the order they were found by the prediction pipeline. For primer design, the homologous flanking regions of the predicted VNTR loci were extracted from all twelve genome sequences and aligned using MUSCLE (<http://www.ebi.ac.uk/Tools/msa/muscle/>; accessed on 27 November 2020) [80]. PCR primers were designed on the conserved regions and tested for the optimal annealing temperature and dimer formation (<http://www.thermoscientific-bio.com/webtools/multipleprimer/>; accessed on 27 November 2020). Specificity of PCR primers and inter-strain size polymorphisms of PCR amplicons were evaluated by in silico PCR (<http://insilico.ehu.es/PCR/>; accessed on 27 November 2020).

## 2.2. Bacterial Strains Sampling and Isolation

The bacterial strains used in this study are listed in Table S1. In order to establish a representative strain collection, a comprehensive survey of pepper plant samples with symptoms of bacterial spot was conducted (Figure 1). Four agro-ecological zones in Bulgaria and three zones in North Macedonia were targeted in this study. Sampling was conducted in diagonal transects in three to four fields at each location. Different parts of affected plants (leaves, stems, flowers, petioles) were collected. For isolation of pathogenic bacteria, 1-cm leaf disks from infected and from healthy plant tissue were cut, disinfected by soaking in a solution of sodium hypochlorite at 2% for 3 min and washed several times with sterile distilled water. Serial ten-fold dilutions to  $10^{-4}$  were performed in sterile physiological salt solution and aliquots of 100  $\mu$ L suspensions were plated onto King's B medium plates. Agar plates were incubated at 28 °C for 48 h. Yellow, mucoid, non-fluorescent colonies typical for *X. euvesicatoria* were observed on all plates coming from symptomatic tissue.



**Figure 1.** Geographic map depicting Europe with a close-up of Bulgaria (top) and North Macedonia (bottom) indicating the origin of the analysed strains, corresponding to four Bulgarian regions (B1 to B4) and three North Macedonian regions (M1 to M3).

### 2.3. Molecular Biological Techniques

A single colony grown on potato-sucrose agar (potatoes 200 g/L, sucrose 20 g/L, and agar 20 g/L) was inoculated in 50 mL lysogenic broth and incubated overnight at 28 °C. The cell density of the bacterial suspension was adapted to an optical density at 600 nm ( $OD_{600nm}$ ) of 1. DNA was extracted from cultivated bacterial cells using the STS kit (STS Ltd., Sofia, Bulgaria) according to the manufacturer's instructions. The yield and purity of the obtained DNA were monitored on a Nanodrop 2000 spectrophotometer (Thermo Scientific).

PCR amplifications of VNTR loci were performed using genomic DNA of *X. euvesicatoria* strains, including three sequenced reference strains as control: 85-10, 66b, and 83M. PCR reactions were carried out in a final volume of 25  $\mu$ L and contained 5  $\mu$ L 5x green GoTaq<sup>®</sup> reaction buffer, 3  $\mu$ L 25 mM MgCl<sub>2</sub>, 1  $\mu$ L oligonucleotide primers at 100  $\mu$ M, 0.5  $\mu$ L 10mM dNTPs, and 0.05  $\mu$ L GoTaq<sup>®</sup> DNA Polymerase (Promega Corp., Madison, WI, USA). Amplification reactions started with an initial denaturation step at 95 °C for 3 min, followed by 35 cycles each consisting of 30 s at 95 °C, 20 s at 55–58 °C (depending on the primer pair, Table S2), and 60 s at 72 °C, and finished by an elongation step of 10 min at 72 °C. PCR-amplified VNTR loci were electrophoretically separated on 1% agarose gels to check for specificity of DNA amplification and sequenced (Beckman Coulter Genomics, UK), using one of the PCR primers (Table S2).

### 2.4. Multilocus Sequence Analysis

Four previously used housekeeping genes were targeted for MLSA: *dnaA*, *fyuA*, *gyrB*, and *rpoD* [20]. DNA amplification was performed as described, and PCR amplicons were sequenced using the forward primers. Sequences, not including the primer sequences, were trimmed to 796 bp (*dnaK*), 628 bp (*gyrB*), 840 bp (*fyuA*), and 846 bp (*rpoD*), totalling to 3110 bp.

### 2.5. VNTR Analysis and Statistics

DNA sequences were aligned using MUSCLE [80] and numbers of complete repeats were derived from multiple sequence alignments. The number of repeats at each locus for every strain was recorded in a matrix and deposited at MLVAbank (<http://www.biopred.net/mlva/>; accessed on 27 November 2020) [62]. Two different datasets were produced. VNTR-rep data represent the alleles scored as repeat numbers whatever the sequence of the repeat is whereas the VNTR-seq dataset separates alleles according to the number of repeats and the sequence of the VNTR locus. The discriminatory power of each locus was evaluated by the Hunter–Gaston Discriminatory Index (HGDI), which was calculated at [http://insilico.ehu.es/mini\\_tools/discriminatory\\_power/](http://insilico.ehu.es/mini_tools/discriminatory_power/) (accessed on 27 November 2020).

The evolutionary relationships among the bacterial strains were displayed as minimum-spanning trees using the software PHYLOViZ version 2.0 [81]. Categorical minimum spanning trees were built using the algorithm recommended for TR data, combining global optimal eBURST (goeBURST) and Euclidean distances [82]. Different estimates of genetic diversity as Nei's index of gene diversity and clonal diversity (Simpson index) were calculated with the poppr 2.2.1 package in R [83]. Allelic richness (A) and private allelic richness (Ap) were computed using the rarefaction procedure for unequal sample sizes with HP-RARE version 1.0 [84]. Nei's index of gene diversity and analysis of molecular variance (AMOVA) were performed using the ARLEQUIN version 3.5 software package [85]. Levels of significance were determined by computing 999 random permutations. Population pairwise  $F_{ST}$  and  $R_{ST}$  using ARLEQUIN were computed for the VNTR-rep dataset and their significance was tested using 999 permutations. The population pairwise  $F_{ST}$  was only computed for the VNTR-seq dataset as alleles are scored as sequence types, which are not expected to evolve under a stepwise mutation model for which  $R_{ST}$  is designed.

## 3. Results

### 3.1. Multilocus Sequence Analysis of Bulgarian and North Macedonian *X. euvesicatoria* Strains

MLSA has been used to determine phylogenetic relationships among bacterial spot agents of pepper and tomato [21,27,86,87]. All these MLSA schemes, and also those that have been developed for the genus *Xanthomonas*, included a portion of the *gyrB* gene [20,88]. We therefore first evaluated the utility of *gyrB* to differentiate *X. euvesicatoria* from the Balkan Peninsula. Based on previous analyses [52,53], three strains from North Macedonia (M) and nine strains from Bulgaria (b) were chosen: 1M, 5M, 7M, 29b, 30b, 38b, 54b, 61b, 67b, 74b, 82b, and 86b. Sequence comparison of a 628-bp DNA fragment revealed that all twelve sequences were identical to each other and to the corresponding sequences from the eight completely sequenced strains of *X. euvesicatoria* pv. *euvesicatoria*, 85-10 (Florida), 83M (North Macedonia), 66b (Bulgaria), and five LMG strains (667, 905, 909, 918, 933) from four continents (India, Ivory Coast, Brazil, Tonga) [13,71,72].

To obtain more sequence information, we further sequenced a portion of the three remaining genes of Young's MLSA scheme, *dnaK*, *fyuA*, and *rpoD* [20]. Again, all DNA sequences from the strains 7M, 29b, 38b, 67b, and 86b, totalling to 3110 bp (including *gyrB*), were identical to each other and to the sequences from the eight reference strains. These data confirm that the *X. euvesicatoria* pepper pathogens are largely monomorphic and that MLSA is not a useful tool for epidemiological studies at any geographical scale.

### 3.2. Identification of Polymorphic VNTR Loci and Development of a 16-Loci MLVA Scheme (MLVA-16)

The screening of twelve available genome sequences identified 76 VNTR loci with 5-bp to 9-bp repeats. Based on conserved presence of the loci and observed polymorphisms, 28 promising loci were selected to develop an MLVA scheme. Candidate VNTR loci were first analysed on a test panel of 15 strains representing worldwide diversity, thus assessing the conservation of the loci (Table S1). Six Bulgarian strains (24b, 27b, 28b, 73b, 89b, and 90b) collected from five different pepper cultivars in two different regions and four North

Macedonian strains (54M, 65M, 66M, and 80M) from two regions and two cultivars were chosen for this purpose. Two pairs of strains isolated from the same cultivar in the same region and in the same year (65M, 66M from North Macedonia and 89b, 90b from Bulgaria) were included to exemplify the resolutive power of the VNTR loci. In addition to the ten strains from the Balkan Peninsula, we included five strains from four more continents. Upon PCR amplification and confirmation by agarose gel electrophoresis, 16 markers were selected for further analyses. All 16 primer pairs led to the amplification of a single dominant DNA fragment, which allowed sequencing using one of the two PCR primers.

Mapping of the 16 loci to the manually annotated genome sequence of strain 85-10, which contains four plasmids, revealed that all of them are encoded on the chromosome without any obvious clustering (Table 1). The majority of repeat arrays are found in intergenic regions. Only two loci, Xe\_03 and Xe\_17, are present in coding sequences. Since both loci consist of 6-bp repeat arrays, repeat number variation does not affect the reading frame. Locus Xe\_17 is found in gene XCV3092, which is predicted to encode an NADPH-dependent sulphite reductase. The repeat array encodes a string of Ala-Asp motifs ([AD]<sub>8</sub> in strain 85-10), which separate the N-terminal flavodoxin domain from the C-terminal oxidoreductase NAD-binding domain. The second locus, Xe\_03, that is found in a coding sequence belongs to the *xopD* type III effector gene, XCV0437, in strain 85-10. Here, the repeat array codes for the motif (Lys-Ala)<sub>3</sub>-(Glu-Ala)<sub>3</sub>-(Lys-Ala)<sub>3</sub>-(Glu-Ala-Lys-Ala)<sub>2</sub>. This motif is found at the end of the so-called N-terminal extension, which had been overlooked during manual annotation but later shown to be translated [89].

**Table 1.** Characteristics of variable number of tandem repeats (VNTR) and Hunter–Gaston Discriminatory Index (HGDI) scores for the 16 loci of the multilocus variable number of tandem repeat analyses (MLVA) scheme.

Locus	Position in Strain 85-10 <sup>1</sup>	Dominant Repeat Type	Other Repeat Types <sup>2</sup>	No. of Alleles <sup>3</sup>	Allelic Range <sup>4</sup>	HGDI Score <sup>5</sup>
Xe_02	215122..215170	TCCCCAT	-	4	4–7 #	0.286
Xe_03	487069..487152	TTTGGC	TCTGGC *	3	12–14 #	0.190
			TTCGGC			
			TTTGTC *			
Xe_04	624229..624277	CGATTCC	-	8	5–12#	0.764
Xe_06	857148..857196	AACAGCC	-	3	6–8 #	0.317
Xe_07	924719..924767	CCGGGTC	CCGGGCC *	4	4–7 #	0.211
Xe_09	1053822..1053863	GGGATTT	GCGGATC	7	6–18	0.803
			GGGAATC			
Xe_10	1222069..1222110	AGGCGGT	AGGCGGC *	6	5–12	0.575
Xe_11	1504314..1504418	CCGATTC	CCTAATC	5 (6)	11–16	0.455 (0.480)
			CCCAATC			
Xe_14	2268785..2268850	ACAGCG	-	6	6–11 #	0.738
Xe_15	3198440..3198527	GCAGACAG	GCAGGCAG	5 (8)	6–10 #	0.688 (0.779)
			GCAGAGAT *			
Xe_16	3505639..3505687	AATGGGG	AATCGGG *	3	5–9	0.263
Xe_17	3514941..3514994	TCGGCA	TCGGCG *	5	9–14	0.437
Xe_22	4396287..4396335	TTGGCGG	TTGGCGC *	3	5–10	0.190
Xe_29	4211581..4211608	CGATTCC	-	2	4–5 #	0.315
			GAATCGG			
Xe_34	458055..458096	GATTCCG	GAATTCG *	4 (5)	5–16	0.190 (0.192)
			GAATCCG *			
Xe_49	4410313..4410342	TGGCCG	-	4	5–8#	0.575
MLVA-16				36		
				haplotypes		0.944

<sup>1</sup> Genomic coordinates in GenBank accession number AM039952. <sup>2</sup> Alternative repeat types that are only found once per repeat array at its end are indicated by a \*. <sup>3</sup> Number of alleles from the VNTR-rep dataset. Numbers in brackets take resolved cases of homoplasy, i.e., different sequence types,

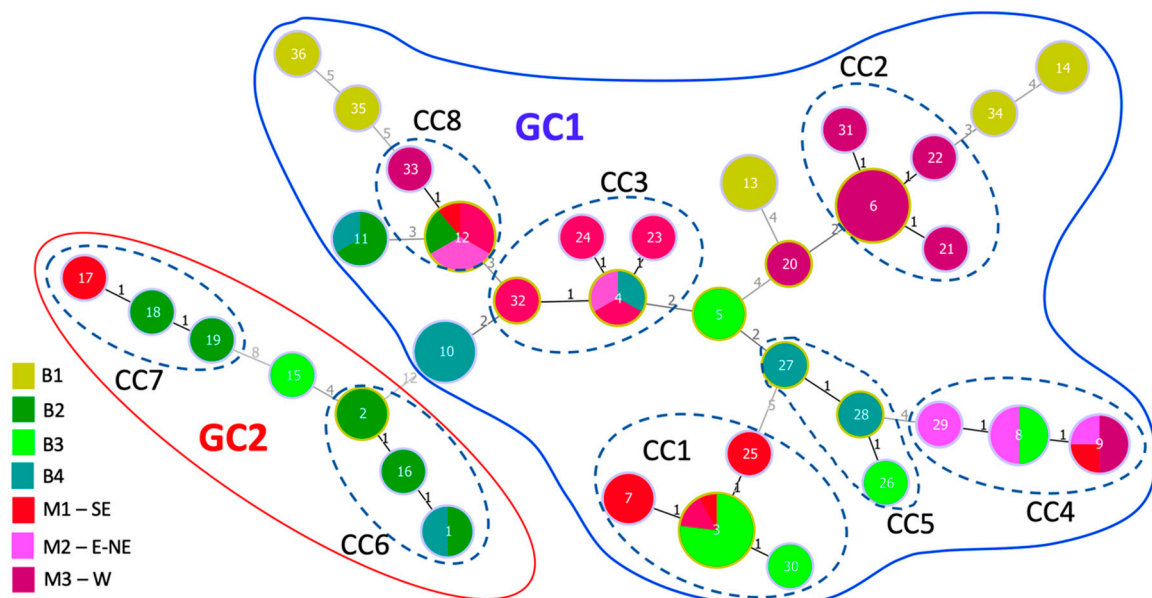


into account (from the VNTR-seq dataset).<sup>4</sup> Allelic ranges that contain all alleles are indicated by a #.<sup>5</sup> Hunter–Gaston discriminatory index (HGDI) scores of individual VNTR loci and of the 16-loci VNTR analysis, MLVA-16. Numbers in brackets correspond to the VNTR-seq dataset.

### 3.3. Application of the MLVA-16 Scheme on a Collection of Pepper-Pathogenic *X. euvesicatoria* from Bulgaria and North Macedonia

Sixteen VNTR loci were used to study a collection of 88 *X. euvesicatoria* strains from Bulgaria and North Macedonia (Table S1). PCR-amplified DNA fragments were sequenced using one of the PCR primers. Numbers of complete repeats were then derived from multiple sequence alignments. The number of alleles per locus ranged from two to eight for loci Xe\_29 and Xe\_04, respectively (Table 1). The number of alleles covered the full allelic range for nine loci out of 16 (Table 1). On average, 4.5 alleles were observed per locus. The discriminant power of the 16 loci, as estimated by HGDI (Hunter–Gaston discriminatory index) scores, ranged from 0.190 (Xe\_03, Xe\_22, and Xe\_34) to 0.803 (Xe\_09). Six loci had poor discriminatory power (HGDI score < 0.3) for the set of *X. euvesicatoria* strains from the Balkan Peninsula (Xe\_02, Xe\_03, Xe\_07, Xe\_16, Xe\_22, Xe\_34). Twenty-five per cent of the VNTR loci had high discriminatory power (HGDI score > 0.6) (Xe\_04, Xe\_09, Xe\_14, Xe\_15).

Combining all 16 loci into an MLVA-16 scheme resulted in an HGDI score of 0.944 for the 88 Balkan strains and allowed to resolve 36 MLVA haplotypes (MTs), thus demonstrating their suitability for typing pepper-pathogenic *X. euvesicatoria* strains. The 36 MTs grouped in eight clonal complexes, i.e., groups of single-locus variants (SLVs), representing 26 MTs and 68 strains (Figure 2). Among these SLVs, 73.7% were single-repeat variants (SRVs) and 21.1% differed by two repeats, which might result from a sampling bias where an evolutionary step is missing. Among the nine loci showing SLVs, seven exhibited exclusively SRVs, Xe\_49 produced a double-locus variant and Xe\_04 evolved with variations involving from one to three repeats.



**Figure 2.** Categorical minimum spanning tree from the VNTR-rep dataset using the goeBURST and Euclidian algorithms. Numbers in circles identify the 36 MLVA haplotypes. The circle sizes are proportional to the number of strains per haplotype, with the smallest circles (e.g., no. 17 in CC7) corresponding to one strain and the largest circle (no. 3 in CC1) corresponding to 13 strains. Colours indicate the country and region of origin, with four Bulgarian regions (B1 to B4) and three North Macedonian regions (SE, South East; E-NE, East-North East; W, West). Black numbers between the circles connect single-locus variants (SLVs) grouped into eight clonal complexes (CC), which are encircled with dashed lines. Gray numbers indicate the number of loci (>1) that differ between connected MLVA haplotypes. Two major genetic clusters, as defined by specific SNPs (GC1 and GC2), are encircled by solid blue (GC1) and red (GC2) lines (see Section 3.5).



### 3.4. Impact of Homoplasmy on VNTR Typing

Since individual repeats within a VNTR locus often vary slightly in sequence, DNA sequencing provides a tool to discover cases of size homoplasmy. This fraction of homoplasious loci is called molecularly accessible size homoplasmy [67]. Indeed, only six loci (Xe\_02, Xe\_04, Xe\_06, Xe\_14, Xe\_29, and Xe\_49) were perfect multiple repeats which consisted of a single repeat type, whereas the other loci were composed of two to four repeat types (Table 1). Almost all the different repeats within composite VNTR loci differed by only one base pair except for locus Xe\_15, where one repeat differed from the other two by two base pairs. However, 12 of the 17 alternative repeat types were found at the end of a repeat array, and only once per repeat array. Only five loci, Xe\_03, Xe\_09, Xe\_11, Xe\_15 and Xe\_34, were really chimeric loci consisting of several alternate repeat types along the entire array.

Among these composite VNTR loci, three (Xe\_11, Xe\_15 and Xe\_34) showed size homoplasmy where the same allele, i.e., number of repeats, could result from different combinations of repeat types or sequence type ST (Table 2). For instance, upon DNA sequencing we noted that locus Xe\_11 consists of three different 7-bp repeat motifs, from which two different combinations produced PCR amplicons with the same size, corresponding to 13 repeats (Table 2). We designated the two homoplasious alleles as Xe\_11-ST1(13) and Xe\_11-ST2(13). While these two alleles could be resolved by DNA sequencing they cannot be resolved by other techniques (e.g., capillary electrophoresis) that only deliver amplicon sizes.

**Table 2.** VNTR loci with size homoplasmy, as revealed by their sequence type (ST).

VNTR Locus	No. of Repeats	ST	Strains <sup>1</sup>	Repeat Pattern
Xe_11	13	1	<u>CC6</u> (62b, 67b, 74b, 96b, 102b)/ <u>CC7</u> (42b, 43b, 35M)/ <u>MT15</u> (13b)	(CCGATTC) <sub>7</sub> -(CCCAATC) <sub>1</sub> -(CCTAATC) <sub>1</sub> -(CCCAATC) <sub>4</sub>
Xe_11	13	2	<u>CC3</u> (10b, 1M, 2M, 5M, 25M, 50M)/ <u>MT10</u> (44b, 45b, 47b, 49b, 51b)	(CCGATTC) <sub>6</sub> -(CCCAATC) <sub>1</sub> -(CCTAATC) <sub>1</sub> -(CCCAATC) <sub>5</sub>
Xe_15	10	1	<u>CC3</u> (10b, 1M, 2M, 5M, 25M, 50M)/ <u>MT11</u> (61b, 69b, 70b)	(GCAGGCAG) <sub>3</sub> -(GCAGACAG) <sub>6</sub> -(GCAGAGAT) <sub>1</sub>
Xe_15	10	2	<u>CC8</u> (105b, 106b, 77M, 79M, 80M, 81M, 82M, 83M, 84M, 86M)/ <u>MT35</u> (80b)	(GCAGGCAG) <sub>2</sub> -(GCAGACAG) <sub>7</sub> -(GCAGAGAT) <sub>1</sub>
Xe_15	9	1	<u>CC1</u> (5b, 12b, 24b, 27b, 29b, 30b, 31b, 38b, 55b, 56b, 7M, 11M, 28M, 31M, 37M, 38M)/ <u>CC5</u> (39b, 93b, 94b)/ <u>MT5</u> (11b, 54b)/ <u>MT10</u> (44b, 45b, 47b, 49b, 51b)/ <u>MT14</u> (81b, 82b)/ <u>MT34</u> (85b)/ <u>MT36</u> (86b)	(GCAGGCAG) <sub>2</sub> -(GCAGACAG) <sub>6</sub> -(GCAGAGAT) <sub>1</sub>
Xe_15	9	2	<u>CC6</u> (62b, 67b, 74b, 96b, 102b)	(GCAGGCAG) <sub>1</sub> -(GCAGACAG) <sub>7</sub> -(GCAGAGAT) <sub>1</sub>
Xe_15	8	1	<u>CC1</u> (28b)/ <u>MT13</u> (77b, 78b, 79b)	(GCAGGCAG) <sub>2</sub> -(GCAGACAG) <sub>5</sub> -(GCAGAGAT) <sub>1</sub>
Xe_15	8	2	<u>CC2</u> (54M, 55M, 56M, 57M, 58M, 61M, 62M, 63M, 64M, 65M, 66M, 67M, 68M, 69M, 70M)/ <u>CC4</u> (89b, 90b, 71M, 72M, 73M, 74M, 76M, 85M, 87M)/ <u>MT20</u> (59M)	(GCAGGCAG) <sub>3</sub> -(GCAGACAG) <sub>4</sub> -(GCAGAGAT) <sub>1</sub>
Xe_34	13	1	<u>CC7</u> (42b, 35M)	(GATTCGG) <sub>5</sub> -(GAATCGG) <sub>2</sub> -(GATTCGG) <sub>1</sub> -(GAATCGG) <sub>4</sub> -(GAATTCG) <sub>1</sub>

Xe_34	13	2	<u>CC6</u> ( <b>62b, 67b, 74b, 96b, 102b</b> )	(GATTCGG) <sub>4</sub> -(GAATCGG) <sub>2</sub> - (GATTCGG) <sub>1</sub> -(GAATCGG) <sub>5</sub> - (GAATTCG) <sub>1</sub>
-------	----	---	--	--

<sup>1</sup> Strains belonging to cluster GC2 are indicated in bold. Clonal complexes (CC) and singleton MLVA haplotypes (MT) are underlined twice if they contain all strains and once if they contain only a subset of strains.

Resolving cases of homoplasmy slightly increased the discriminatory power of the three affected loci (Table 1). Yet, correcting for homoplasmy did not change the number of MLVA haplotypes and likewise did not increase the clonal diversity estimated by the Simpson index (Table 3).

**Table 3.** Genetic diversity parameters of *X. euvesicatoria* from Bulgaria ( $n = 45$ ) and North Macedonia ( $n = 43$ ) estimated from the two MLVA-16 datasets (VNTR-rep and VNTR-seq).

Country	Polymorphic Loci	eMLG <sup>1</sup>	Simpson Index D	He <sup>2</sup> (seq) <sup>4</sup>	A <sup>3</sup> (seq) <sup>4</sup>	Ap <sup>3</sup> (seq) <sup>4</sup>
Both countries		22.9	0.944	0.437	-	-
Bulgaria	16	21.5	0.932	0.494 (0.506)	4.25 (4.56)	1.25 (1.44)
North Macedonia	15 <sup>5</sup>	18	0.890	0.329 (0.332)	3.13 (3.25)	0.13 (0.13)

<sup>1</sup> eMLG, expected number of multilocus genotypes estimated from a rarefaction procedure ( $n = 43$ ).

<sup>2</sup> He, Nei's index of gene diversity. <sup>3</sup> A, allelic richness; and Ap, private allelic richness estimated from a rarefaction procedure ( $n = 43$ ). <sup>4</sup> Numbers are estimated from the VNTR-rep dataset whereas numbers in brackets are from the VNTR-seq dataset, which take resolved cases of homoplasmy into account (homoplasmy for VNTR loci Xe\_11, Xe\_15 and Xe\_34). <sup>5</sup> The VNTR locus Xe\_02 is monomorphic for North Macedonian strains.

### 3.5. Single-Nucleotide Polymorphisms in the Flanking Sequences

DNA sequencing did not only reveal sequences of the tandem repeats but also provided insight into the sequence conservation of the flanking regions of the repeat arrays. Similar to the four housekeeping genes, DNA sequences were highly conserved. Yet, two VNTR loci showed single nucleotide polymorphisms (SNPs) in their flanking sequences. For ten strains (13b, 42b, 43b, 62b, 67b, 73b, 74b, 96b, 102b, 35M), we observed a C to G exchange 26 base pairs upstream of Xe\_09 within the coding sequence of gene XCV0924, leading to a silent mutation (threonine codon ACC → ACG). Two additional SNPs were shared among the same ten strains in the intergenic region between *fliP* (XCV1988 in strain 85-10) and *fliQ* (XCV1987 in strain 85-10), 4 and 119 bp upstream of Xe\_14. Hence, information from the conserved flanking regions divided the set of strains into two genetic clusters (GC), one consisting of ten strains (nine from Bulgaria and one from North Macedonia) sharing three SNPs, which comprise seven MTs (1, 2, 15, 16, 17, 18, 19) (GC2), and another cluster comprising the rest of the strains (GC1; Figure S1).

Interestingly, all ten strains of GC2 shared the same flanking sequences of Xe\_09 and Xe\_14 with strains 85-10 (Florida), 66b (Bulgaria), and four LMG strains (667, 905, 909, 933 from New Zealand, India, Ivory Coast, Brazil), collectively originating from six continents, but not with strain 83M from North Macedonia, which belongs to GC1 (Table S1) [13,71,72]. This finding suggests that strains of GC2 are phylogenetically linked to a worldwide lineage.

### 3.6. Population Structure of *X. euvesicatoria*

The population of *X. euvesicatoria* from Bulgaria showed a slightly greater genetic diversity than the North Macedonian population, as shown by the expected number of multilocus genotypes (eMLG = 21.5 and 18.0, respectively) and the allelic richness, with A = 4.25 and 3.13, respectively (Table 3). The private allelic richness was much greater for the Bulgarian population (Ap = 1.25) than for the North Macedonian strains (Ap = 0.13). These estimates of genetic diversity based on allelic data were a little higher when

estimated from the VNTR-seq dataset (Table 3, numbers in brackets) but those estimated from haplotypes did not change.

The Bulgarian and North Macedonian populations were clearly differentiated when computed from both the VNTR-rep dataset ( $F_{ST} = 0.101$ ,  $p < 0.001$ / $R_{ST} = 0.108$ ,  $p = 0.002$ ) and from the sequence-corrected VNTR-seq dataset (with  $F_{ST} = 0.098$ ,  $p < 0.001$ ).

Analysis of molecular variance (AMOVA) partitioning different regional levels revealed that most of the genetic variation was significantly explained by the variation between regions within countries and by the intra-region variations. In contrast, inter-country variation did not significantly contribute to the total genetic variation (1.05%,  $p = 0.47$ ) (Table 4). Most of the pairwise comparisons of the regional populations were significantly differentiated (Table 5). However, none of the pairwise  $R_{ST}$  parameters between Bulgarian regions B3 and B4, the North Macedonian region M1 and the North Macedonian region M2 bordering Bulgaria were significantly different (Table 5).

The haplotypic diversity estimates (eMLG and Simpson index) of Bulgarian and North Macedonian regional populations were in the same range (Table S3). However, the allelic diversity estimates were slightly greater for the Bulgarian regional populations, which had also more private alleles than the North Macedonian regions.

**Table 4.** Analysis of molecular variance (AMOVA) partitioning different geographical levels (regions and countries) estimated from  $R_{ST}$  parameters based on the VNTR-rep dataset.

Source of Variation	D.f.	Sum of Squares	Variance Components	Percentage of Variation	<i>p</i> -Value
- between countries	1	177.11	0.329	1.05	0.437
- between regions within countries	5	706.15	10.05	32.14	<0.001
- within regions	81	1691.68	20.88	66.80	<0.001
Total	87	2574.94	31.26		

**Table 5.** Genetic differentiation estimated by  $R_{ST}$  pairwise comparisons of regional collections of *X. euvesicatoria* strains from Bulgaria (B1-B4) and North Macedonia (M1-M3) based on the VNTR-rep dataset.

	B2	B3	B4	M1	M2	M3
B1	0.442 **	0.304 ***	0.239 **	0.250 **	0.321 *	0.501 ***
B2		0.400 **	0.358 **	0.406 **	0.447 *	0.645 ***
B3			0.012 NS	−0.031 NS	0.070 NS	0.313 ***
B4				−0.040 NS	−0.014 NS	0.200 **
M1					−0.001 NS	0.237 ***
M2						0.451 ***

*p* values: NS, non significant, i.e.,  $p > 0.05$ ; \*,  $0.01 < p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ .

### 3.7. Strains from Bulgaria and North Macedonia Are Genetically Closely Related

The relationship among the strains based on the MLVA-16 results is presented in a minimum-spanning tree (Figure 2). The minimum spanning tree produced from the VNTR-rep dataset grouped the 36 MLVA haplotypes (MTs) in eight small clonal complexes (68 strains) and ten singletons (20 strains), i.e., haplotypes that differed by more than one locus from all other haplotypes. The clonal complexes are formed by two (CC8) to four (CC1, CC2, CC3) haplotypes. Five clonal complexes (CC1, CC3, CC4, CC7, CC8) grouped strains from both countries, two other clonal complexes (CC5 and CC6) grouped strains from different regions of Bulgaria, and the last clonal complex (CC2) grouped only strains from the Western region of North Macedonia (Figure 2). A sub-group of nine MTs including CC3, CC5 and MTs 5 and 10 are double-locus variants, perhaps due to a sampling bias with an evolutionary step missing, and could thus be epidemiologically related.

Strains from the different regions of Bulgaria and North Macedonia were found to be widespread in genetically distant clonal complexes or singletons. Several strains of

different regions, except for B1, shared haplotypes with strains from other regions or the other country (Figure 2). The large geographical distribution of haplotypes and the existence of epidemiologically closely related strains in different regions and countries suggest long dispersal of strains.

Most of the haplotypes of the network differed by five loci or less and grouped as a large genetic cluster, corresponding to GC1. This cluster was well separated from a phylogenetically distant second cluster (GC2) composed of two clonal complexes (CC6, CC7) and the singleton MT15 by variations at 12 loci or more. Notably, cluster GC2 contained mostly strains from three Bulgarian regions and one strain from North Macedonia.

Interestingly, only the GC2 strains showed the SNPs detected in the flanking regions of loci Xe\_09 and Xe\_14 (see before), in comparison to GC1 strains. In addition to these two polymorphic loci, specific alleles of GC2 were observed for eight loci (Xe\_03, Xe\_07, Xe\_10, Xe\_11, Xe\_15, Xe\_17, Xe\_22, Xe\_34). Among them, Xe\_34, which was monomorphic in GC1 strains sharing a unique allele of four repeats, had different haplotypes (12 to 15 repeats) and two homoplasious alleles with 13 repeats in GC2 strains. Similarly, GC2 strains had specific sequence types for two additional homoplasious alleles, i.e., ST1 for Xe\_11 (all strains) and ST2 for the Xe\_15 allele with nine repeats (CC6) (Table 2).

The relationships among the haplotypes in the minimum spanning tree obtained from sequence-corrected data (VNTR-seq dataset) did not change much except that cluster GC2 strains were placed as a 13-loci variant of MT8 (CC4) instead of a 12-loci variant of MT10 (Figure S2). However, considering the large distance between GC1 and GC2 it is not very meaningful to connect them at all. Rather, our data suggested that our sample consists of two groups of strains with one of them (GC2) being linked to a worldwide pandemic expansion and another one (GC1) being restricted to the Balkan Peninsula at the time being.

### 3.8. Transferability of VNTR Markers to Other Pathovars of *X. euvesicatoria*

Here, we used genomic information for strains from different pathovars of *X. euvesicatoria*, including *perforans* (strain 91–118), *alfalfae* (strain CFBP 3836), *allii* (strain CFBP 6369) and *citrumelo* (strain F1), for prediction of VNTRs and primer design. This way, we expected to develop a scheme that can be used beyond the pathovar *euvesicatoria*. To test this hypothesis, we extracted the 16 VNTR loci from all available genome sequences of *X. euvesicatoria* and analysed them in silico (Table 6). Except for Xe\_03, 15 out of the 16 loci were present in all the other pathovars (*alfalfae*, *allii*, *citrumelo*, *commiphorea*, *dieffenbachiae*, *perforans*). Interestingly, pathovar *perforans*, which causes the same disease as the pathovar *euvesicatoria* but was ranked as a different species when we started the project, shared all 16 loci with pathovar *euvesicatoria*. Remarkably, most loci in pathovar *perforans* were polymorphic and only two of them, Xe\_06 and Xe\_07, appeared to be monomorphic. Compared with pathovar *euvesicatoria*, allelic ranges in *perforans* appeared to be smaller, but this observation may be due to sampling biases, as most *perforans* strains were from a restricted geographic area.

**Table 6.** Allelic range of VNTR loci from different pathovars of *Xanthomonas euvesicatoria*.

Locus	<i>euvesicatoria</i>	<i>perforans</i>	<i>alfalfae</i>	<i>allii</i>	<i>citrumelo</i>	<i>commiphoreae</i>	<i>dieffenbachiae</i>
Xe_02	2–11	1–16	13	10	12	18	15
Xe_03	8–14	11–12	NA	NA	NA	NA	NA
Xe_04	3–11	2–3	3	9	3–4	4	4
Xe_06	3–11	2	2	4	2	6	2
Xe_07	3–8	4	2	4	2–8	6	6
Xe_09	5–16	5–7 (11)	5–6	12	12–13	6	14
Xe_10	3–8	(3) 5–7	4	4	4	5	10
Xe_11	9–17	4–6	4	9	5–6	9	4
Xe_14	2–13	5–8	5	9	8–9	6	9
Xe_15	6–11	2 (6.5)	2	2	4–6	6	2

Xe_16	3–16	(4) 10–12	12	4	4	4	11
Xe_17	8–18	10–16	9	15	9–12	15	10
Xe_22	5–21	3–9	5	12	4	16	8
Xe_29	3–4	4–7	2–9	5	5	4	3
Xe_34	4–15	9–14	7–12	8	8–11	15	7
Xe_49	7–18	(4) 6–7	6–7	4	9–15	10	9
No. of strains	54	143	2	1	4	1	1

#### 4. Discussion

Here, we developed and applied the first VNTR scheme for the pepper- and tomato-pathogenic strains of *X. euvesicatoria*. VNTR typing has many advantages in comparison with other techniques with a similar level of taxonomic resolution, such as amplified fragment length polymorphism (AFLP), randomly amplified polymorphic DNA (RAPD) and repetitive sequence-based rep-PCR [90]. Whereas traditional fingerprinting techniques generate more or less complex profiles, which are often technically demanding and hard –if not impossible– to compare between laboratories, VNTRs deliver integer numbers of repeats, which can be easily electronically stored and compared. VNTRs represent single-locus markers that are extremely informative and can be analyzed as allele frequencies for population genetics and for effectively tracing bacterial outbreaks or dispersal. This is also the reason why MLVA schemes, i.e., VNTR schemes targeting between 8 and >20 molecular markers, have become more and more popular in the field of plant pathology. Strikingly, MLVA schemes are now available for seven economically important species of *Xanthomonas*, namely *X. arboricola* (pvs. *corylina*/common hazel, *fragariae*/strawberry, *juglandis*/Persian walnut, *populi*/poplar and other trees, *pruni*/stone fruits like almond, apricot, cherry, peach and plum), *X. citri* (pvs. *citri*/citrus, *mangiferaeindicae*/mango, *viticola*/grapevine), *X. euvesicatoria* (pv. *euvesicatoria*/pepper and tomato), *X. fragariae* (strawberry)], *X. oryzae* (pvs. *oryzae* and *oryzicola*)/rice), *X. phaseoli* (pv. *manihotis*/cassava) and *X. vasicola* (pv. *musacearum*/banana) [55,56,58,60–65,91–95]. In order to allow and stimulate comparisons with other datasets, we deposited our data at MLVAbank (<http://www.biopred.net/MLVA/>, accessed on 27 November 2020).

Whereas existing MLSA schemes target all members of the *Xanthomonas* genus, they suffer from a comparatively low level of taxonomic resolution and cannot efficiently resolve strains within a pathovar [20,92,96]. On the other hand, high-resolution MLVA schemes targeting small repeats, i.e., microsatellites, are typically developed for a distinct pathovar and are limited in their applicability to other pathovars of the same species. For instance, only a subset of the VNTR loci that were employed for an MLVA scheme of *X. oryzae* pv. *oryzae* could also be used for the other pathovar, *oryzicola* [58,62]. Similarly, starting with a set of 26 VNTR loci for the species *X. arboricola*, only 9 to 23 of them were useful to characterize strains from five different pathovars [60]. Yet, in some cases, schemes developed for a distinct pathovar proved to be useful for other pathovars of the same species [55,56]. Based on available genome sequences we estimated the transferability of our VNTR markers to other pathovars of *X. euvesicatoria*. In total, 14 of the 16 VNTR markers were found to be polymorphic in the pathovar *perforans* and thus useful for molecular typing. We are not sure how valuable, i.e., polymorphic, the loci are for the other pathovars because we had only access to very few genome sequences for them. However, at least for those VNTRs with moderate or large repeat numbers (i.e., four or more), we expect size polymorphisms for most of them [97].

Size homoplasy originating from different combinations of repeats was detected for five haplotypes in three of 16 VNTR loci (Xe\_11, Xe\_15, Xe\_34; Table 2), as revealed by DNA sequencing. In these cases, haplotypes with the same number of repeats resulted from different combinations of repeat types in these loci. Although size homoplasy at three VNTR loci somewhat reduced the genetic diversity score, this effect did not have a major impact on the population structure and phylogenetic relationships among the

analyzed strains. Sequence variation was also observed in the flanking regions of two loci Xe\_09 and Xe\_14 but they did not lead to size homoplasy. Remarkably, sequence types were lineage-dependent, thus supporting the population structure with two major lineages. Moreover, size homoplasy did not occur within clonal complexes but between genetically remote strains that could originate from different populations as shown previously for different populations of *X. citri* pv. *citri* [93]. Therefore, the information from the sequences (VNTR-seq datasets) strongly supports the conclusion drawn from alleles scored as repeat numbers (VNTR-rep datasets).

While DNA sequencing allowed to decipher the relatively small impact of homoplasy on typing, it may not be advantageous and cost-saving to use this approach for routine typing. However, this MLVA-16 scheme could easily be used in a high-throughput typing technique based on capillary sequencer technology and a multiplexing strategy [96]. Primer pooling is based on annealing temperature (Table S2), expected amplicon size and the use of different fluorescent dyes. We propose to multiplex four primer pairs, each with a specific dye (Figure S3), as previously reported for other xanthomonads [62,65,95].

The sequence polymorphisms in the flanking regions of two VNTR loci (Xe\_09, Xe\_14), i.e., three SNPs, divided the sample into two major lineages, GC1 and GC2, which correlated with the gross population structure based on the MLVA haplotypes. We therefore conclude that the minimum spanning tree with its two major branches is robust because this topology is supported by the three SNPs, which are likely to evolve slower. Interestingly, SNPs shared among GC2 strains were also observed in the genome sequences of other strains originating from six continents (Europe, North America, South America, Africa, Asia, Oceania). This finding suggests that strains of this group are phylogenetically linked to a worldwide lineage and prompted us to check if sequence types that appeared to be restricted to the Balkan Peninsula have been found elsewhere. BLASTN analysis of whole-genome sequences of pepper/tomato pathogens at NCBI GenBank revealed that 46 *X. euvesicatoria* strains share the three SNPs with GC1 strains whereas only five strains contain the three SNPs of GC2 strains, among them 83M from North Macedonia, strains 259, 315 and 329 from the United States of America, and strain BRIP62438 from Australia. Three strains at GenBank did not share any of the sequence types at the flanking region of Xe\_14, among them an atypical strain (LMG 918), which does not cluster with any of the two pepper/tomato-pathogenic pathovars of *X. euvesicatoria* (pvs. *euvesicatoria* or *perforans*) [13,98], strain NI38, which is another atypical strain originating from Nigeria [86], and strain LMG 27970, for which the assembled genome sequence with its 1491 contigs lacks this locus [14].

Molecular epidemiology is based on our ability to differentiate the individuals and to estimate their genetic relatedness to decipher the possible propagative pathways and the origin of the inoculum. Both the high discriminatory power of the VNTR markers and the sequence variation within these loci supported the existence of two major lineages of *X. euvesicatoria* pv. *euvesicatoria*, one of which appears to be overrepresented on the Balkan Peninsula. The smaller GC2 corresponds to worldwide distributed SNP haplotypes in the flanking region of two VNTR loci, suggesting introduction into Bulgaria and North Macedonia from a dominant worldwide lineage, whereas the majority of strains (GC1), distributed in all the sampled regions of Bulgaria and North Macedonia, are endemic. Our data support at least three introduction events at the origin of the Balkan population of *X. euvesicatoria*. The two clonal complexes of GC2 are phylogenetically remote and could originate from two different introduction events. GC1 represents a group of more or less genetically related strains whose ancestor was probably introduced during an older event. Our data further indicate a strong genetic link between the North Macedonian and Bulgarian populations of *X. euvesicatoria*. A part of the North Macedonian population probably originated from Bulgaria, consecutively to one or few introduction events. This conclusion is supported by the date of disease reports, greater genetic diversity and private allelic richness observed in Bulgaria, the absence of genetic differentiation between the



North Macedonian regions bordering Bulgaria and some Bulgarian regions and the sharing of haplotypes or clonal complexes by strains from those regions.

Long-distance dispersal through infected plant material (seeds) most probably occurred within and between countries. Seeds can serve as an important source of inoculum and this could favor the introduction of new strains [4]. More analyses of strains from various regions in the world are required to definitively disclose the origin of these two lineages of *X. euvesicatoria* in the Balkan area. Our MLVA tool together with the informative SNPs will be instrumental in such analyses and will contribute to a better understanding of dispersal pathways and the role of infected seeds in the epidemiology of *X. euvesicatoria*.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2076-2607/9/3/536/s1](http://www.mdpi.com/2076-2607/9/3/536/s1), Figure S1: Categorical minimum spanning tree from the VNTR-rep dataset using the goeBURST and Euclidian algorithms, Figure S2: Categorical minimum spanning tree from the VNTR-seq dataset using the goeBURST algorithm, Figure S3: Proposed VNTR multiplexing scheme for pools of four PCR primer pairs, Table S1: Bacterial strains used in this study and their meta and VNTR typing data, Table S2: Oligonucleotide primers used in this study, Table S3: Genetic diversity parameters estimated for *X. euvesicatoria* strains from different regions of Bulgaria and North Macedonia.

**Author Contributions:** Conceptualization, R.K., T.V. and C.V.; methodology, R.K., T.V. and C.V.; validation, R.K., T.V. and C.V.; formal analysis, R.K., T.V. and C.V.; resources, N.B., P.M., S.M. and T.V.; data curation, R.K. and C.V.; writing—original draft preparation, R.K., T.V. and C.V.; writing—review and editing, N.B., R.K., P.M., S.M., T.V. and C.V.; visualization, C.V.; supervision, R.K.; project administration, R.K.; funding acquisition, R.K. and T.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Agence National de la Recherche (ANR), France, grant number ANR-10-BLAN-1723, and by the Bulgarian Ministry of Education, Youth and Science, grant number BG051PO001-3.3.05-0001. The APC was funded by COST (European Cooperation in Science and Technology).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data produced in this study are presented in this publication.

**Acknowledgments:** This article is based upon work from COST Action CA16107 EuroXanth, sup.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Jones, J.B.; Stall, R.E.; Bouzar, H. Diversity among xanthomonads pathogenic on pepper and tomato. *Annu. Rev. Phytopathol.* **1998**, *36*, 41–58, doi:10.1146/annurev.phyto.36.1.41.
2. Stall, R.E.; Jones, J.B.; Minsavage, G.V. Durability of resistance in tomato and pepper to xanthomonads causing bacterial spot. *Annu. Rev. Phytopathol.* **2009**, *47*, 265–284, doi:10.1146/annurev-phyto-080508-081752.
3. European and Mediterranean Plant Protection Organization. *Xanthomonas* spp. (*Xanthomonas euvesicatoria*, *Xanthomonas gardneri*, *Xanthomonas perforans*, *Xanthomonas vesicatoria*) causing bacterial spot of tomato and sweet pepper. *EPPO Bull.* **2013**, *43*, 7–20, doi:10.1111/epp.12018.
4. Potnis, N.; Timilsina, S.; Strayer, A.; Shantheraj, D.; Barak, J.D.; Paret, M.L.; Vallad, G.E.; Jones, J.B. Bacterial spot of tomato and pepper: diverse *Xanthomonas* species with a wide variety of virulence factors posing a worldwide challenge. *Mol. Plant Pathol.* **2015**, *16*, 907–920, doi:10.1111/mpp.12244.
5. Ritchie, D.F. Bacterial spot of pepper and tomato. In *The Plant Health Instructor*; American Phytopathology Society: St Paul, MN, USA, 2000; doi:10.1094/PHI-I-2000-1027-01.
6. Stall, R.E.; Beaulieu, C.; Egel, D.; Hodge, N.C.; Leite, R.P.; Minsavage, G.V.; Bouzar, H.; Jones, J.B.; Alvarez, A.M.; Benedict, A.A. Two genetically diverse groups of strains are included in *Xanthomonas campestris* pv. *vesicatoria*. *Int. J. Syst. Bacteriol.* **1994**, *44*, 47–53, doi:10.1099/00207713-44-1-47.
7. Dowson, W. On the systematic position and generic names of the Gram-negative bacterial plant pathogens. *Zentralbl. Bakteriol. Parasitenk. Infekt. Hyg. Abt. II* **1939**, *100*, 177–193.

8. Vauterin, L.; Hoste, B.; Kersters, K.; Swings, J. Reclassification of *Xanthomonas*. *Int. J. Syst. Bacteriol.* **1995**, *45*, 472–489, doi:10.1099/00207713-45-3-472.
9. Šutić, D. Tomato bacteriosis. *Rev. Appl. Mycol.* **1957**, *36*, 734–735.
10. Jones, J.B.; Bouzar, H.; Stall, R.E.; Almira, E.C.; Roberts, P.D.; Bowen, B.W.; Sudberry, J.; Strickler, P.M.; Chun, J. Systematic analysis of xanthomonads (*Xanthomonas* spp.) associated with pepper and tomato lesions. *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 1211–1219, doi:10.1099/00207713-50-3-1211.
11. Dye, D. Cultural and biochemical reaction of additional *Xanthomonas* species. *N. Z. J. Sci.* **1966**, *9*, 913–919.
12. Jones, J.B.; Lacy, G.H.; Bouzar, H.; Stall, R.E.; Schaad, N.W. Reclassification of the xanthomonads associated with bacterial spot disease of tomato and pepper. *Syst. Appl. Microbiol.* **2004**, *27*, 755–762, doi:10.1078/0723202042369884.
13. Barak, J.D.; Vancheva, T.; Lefeuvre, P.; Jones, J.B.; Timilsina, S.; Minsavage, G.V.; Vallad, G.E.; Koebnik, R. Whole-genome sequences of *Xanthomonas euvesicatoria* strains clarify taxonomy and reveal a stepwise erosion of type 3 effectors. *Front. Plant Sci.* **2016**, *7*, 1805, doi:10.3389/fpls.2016.01805.
14. Constantin, E.C.; Cleenwerck, I.; Maes, M.; Baeyen, S.; Van Malderghem, C.; De Vos, P.; Cottyn, B. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant Pathol.* **2016**, *65*, 792–806, doi:10.1111/ppa.12461.
15. Oren, A.; Garrity, G.M. Notification of changes in taxonomic opinion previously published outside the IJSEM. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 2081–2086, doi:10.1099/ijsem.0.002071.
16. Timilsina, S.; Kara, S.; Jacques, M.-A.; Potnis, N.; Minsavage, G.V.; Vallad, G.E.; Jones, J.B.; Fischer-Le Saux, M. Reclassification of *Xanthomonas gardneri* (ex Šutić 1957) Jones et al., 2006 as a later heterotypic synonym of *Xanthomonas cynarae* Trébaol et al., 2000 and description of *X. cynarae* pv. *cynarae* and *X. cynarae* pv. *gardneri* based on whole genome analyses. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 343–349, doi:10.1099/ijsem.0.003104.
17. Morinière, L.; Burlet, A.; Rosenthal, E.R.; Nesme, X.; Portier, P.; Bull, C.T.; Lavire, C.; Fischer-Le Saux, M.; Bertolla, F. Clarifying the taxonomy of the causal agent of bacterial leaf spot of lettuce through a polyphasic approach reveals that *Xanthomonas cynarae* Trébaol et al., 2000 emend. Timilsina et al., 2019 is a later heterotypic synonym of *Xanthomonas hortorum* Vauterin et al. 1995. *Syst. Appl. Microbiol.* **2020**, *43*, 126087, doi:10.1016/j.syapm.2020.126087.
18. Parkinson, N.; Aritua, V.; Heeney, J.; Cowie, C.; Bew, J.; Stead, D. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 2881–2887, doi:10.1099/ijms.0.65220-0.
19. Parkinson, N.; Cowie, C.; Heeney, J.; Stead, D. Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences. *Int. J. Syst. Evol. Microbiol.* **2009**, *59*, 264–274, doi:10.1099/ijms.0.65825-0.
20. Young, J.; Park, D.-C.; Shearman, H.; Fargier, E. A multilocus sequence analysis of the genus *Xanthomonas*. *Syst. Appl. Microbiol.* **2008**, *31*, 366–377, doi:10.1016/j.syapm.2008.06.004.
21. Hamza, A.A.; Robène-Soustrade, I.; Jouen, E.; Lefeuvre, P.; Chiroleu, F.; Fischer-Le Saux, M.; Gagnevin, L.; Pruvost, O. Multi-Locus Sequence Analysis- and Amplified Fragment Length Polymorphism-based characterization of xanthomonads associated with bacterial spot of tomato and pepper and their relatedness to *Xanthomonas* species. *Syst. Appl. Microbiol.* **2012**, *35*, 183–190, doi:10.1016/j.syapm.2011.12.005.
22. Schwartz, A.R.; Potnis, N.; Timilsina, S.; Wilson, M.; Patané, J.; Martins, J.J.; Minsavage, G.V.; Dahlbeck, D.; Akhunova, A.; Almeida, N.; et al. Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. *Front. Microbiol.* **2015**, *6*, 535, doi:10.3389/fmicb.2015.00535.
23. Bouzar, H.; Jones, J.; Somodi, G.; Stall, R.; Daouzli, N.; Lambe, R.; Gastelum, R.F.; Correa, R.T. Diversity of *Xanthomonas campestris* pv. *vesicatoria* in tomato and pepper fields of Mexico. *Can. J. Plant Pathol.* **1996**, *18*, 75–77, doi:10.1080/07060669609500659.
24. Bouzar, H.; Jones, J.B.; Stall, R.E.; Louws, F.J.; Schneider, M.; Rademaker, J.L.W.; De Bruijn, F.J.; Jackson, L.E. Multiphasic analysis of xanthomonads causing bacterial spot disease on tomato and pepper in the Caribbean and Central America: Evidence for common lineages within and between countries. *Phytopathology* **1999**, *89*, 328–335, doi:10.1094/phyto.1999.89.4.328.
25. Quezado-Duval, A.M.; Leite, R.P., Jr.; Truffi, D.; Camargo, L.E.A. Outbreaks of bacterial spot caused by *Xanthomonas gardneri* on processing tomato in Central-West Brazil. *Plant Dis.* **2004**, *88*, 157–161, doi:10.1094/pdis.2004.88.2.157.
26. Cuppels, D.A.; Louws, F.J.; Ainsworth, T. Development and evaluation of PCR-based diagnostic assays for the bacterial speck and bacterial spot pathogens of tomato. *Plant Dis.* **2006**, *90*, 451–458, doi:10.1094/pd-90-0451.
27. Hamza, A.A.; Robène-Soustrade, I.; Jouen, E.; Gagnevin, L.; Lefeuvre, P.; Chiroleu, F.; Pruvost, O. Genetic and pathological diversity among *Xanthomonas* strains responsible for bacterial spot on tomato and pepper in the Southwest Indian Ocean region. *Plant Dis.* **2010**, *94*, 993–999, doi:10.1094/pdis-94-8-0993.
28. Ma, X.; Lewis Ivey, M.L.; Miller, S.A. First report of *Xanthomonas gardneri* causing bacterial spot of tomato in Ohio and Michigan. *Plant Dis.* **2011**, *95*, 1584, doi:10.1094/PDIS-05-11-0448.
29. Kebede, M.; Timilsina, S.; Ayalew, A.; Admassu, B.; Potnis, N.; Minsavage, G.V.; Goss, E.M.; Hong, J.C.; Strayer, A.; Paret, M.; et al. Molecular characterization of *Xanthomonas* strains responsible for bacterial spot of tomato in Ethiopia. *Eur. J. Plant Pathol.* **2014**, *140*, 677–688, doi:10.1007/s10658-014-0497-3.
30. Araújo, E.; Costa, J.; Ferreira, M.; Quezado-Duval, A. Widespread distribution of *Xanthomonas perforans* and limited presence of *X. gardneri* in Brazil. *Plant Pathol.* **2016**, *66*, 159–168, doi:10.1111/ppa.12543.
31. Osdaghi, E.; Taghavi, S.M.; Hamzehzarghani, H.; Fazliarab, A.; Lamichhane, J.R. Monitoring the occurrence of tomato bacterial spot and range of the causal agent *Xanthomonas perforans* in Iran. *Plant Pathol.* **2017**, *66*, 990–1002, doi:10.1111/ppa.12642.

32. Osdaghi, E.; Taghavi, S.M.; Koebnik, R.; Jibrin, M.O.; Lamichhane, J.R. Multilocus sequence analysis reveals a novel phylogroup of *Xanthomonas euvesicatoria* pv. *perforans* causing bacterial spot of tomato in Iran. *Plant Pathol.* **2018**, *67*, 1601–1611, doi:10.1111/ppa.12864.
33. Jones, J.B.; Pohronezny, K.L.; Stall, R.E.; Jones, J.P. Survival of *Xanthomonas campestris* pv. *vesicatoria* in Florida on tomato crop residue, weeds, seeds, and volunteer tomato plants. *Phytopathology* **1986**, *76*, 430–434, doi:10.1094/phyto-76-430.
34. Jones, J.B.; Stall, R.E.; Scott, J.W.; Somodi, G.C.; Bouzar, H.; Hodge, N.C. A third tomato race of *Xanthomonas campestris* pv. *vesicatoria*. *Plant Dis.* **1995**, *79*, 395–398, doi:10.1094/pd-79-0395.
35. Newberry, E.A.; Bhandari, R.; Minsavage, G.V.; Timilsina, S.; Jibrin, M.O.; Kemble, J.; Sikora, E.J.; Jones, J.B.; Potnis, N. Independent evolution with the gene flux originating from multiple *Xanthomonas* species explains genomic heterogeneity in *Xanthomonas perforans*. *Appl. Environ. Microbiol.* **2019**, *85*, 00885–19, doi:10.1128/AEM.00885-19.
36. Timilsina, S.; Pereira-Martin, J.A.; Minsavage, G.V.; Iruegas-Bocardo, F.; Abrahamian, P.; Potnis, N.; Kolaczowski, B.; Vallad, G.E.; Goss, E.M.; Jones, J.B. Multiple recombination events drive the current genetic structure of *Xanthomonas perforans* in Florida. *Front. Microbiol.* **2019**, *10*, 448, doi:10.3389/fmicb.2019.00448.
37. EFSA Panel on Plant Health. Scientific opinion on the pest categorisation of *Xanthomonas campestris* pv. *vesicatoria* (Doidge) Dye. *EFSA J.* **2014**, *12*, 3720, doi:10.2903/j.efsa.2014.3720.
38. Kizheva, Y.; Vancheva, T.; Hristova, P.; Stoyanova, M.; Stojanovska, M.; Moncheva, P.; Bogatzevska, N. Identification of *Xanthomonas* strains from tomato and pepper and their sensitivity to antibiotics and copper. *Bulg. J. Agric. Sci.* **2013**, *19*, 80–82.
39. EPPO A2 List of Pests Recommended for Regulation as Quarantine Pests. Available online: [https://www.eppo.int/ACTIVITIES/plant\\_quarantine/A2\\_list](https://www.eppo.int/ACTIVITIES/plant_quarantine/A2_list) (accessed on 27 November 2020).
40. Buonauro, R.; Stravato, V.M.; Scortichini, M. Characterization of *Xanthomonas campestris* pv. *vesicatoria* from *Capsicum annuum* L. in Italy. *Plant Dis.* **1994**, *78*, 296–299, doi:10.1094/pd-78-0296.
41. Zaccardelli, M.; Campanile, F.; Villecco, D.; Parisi, M. Infections of bacterial spot on processing tomato in southern Italy. *ISHS Acta Hort.* **2011**, *914*, 71–73, doi:10.17660/actahortic.2011.914.10.
42. Pernezny, K.; Kúdela, V.; Kokošková, B.; Hládká, I. Bacterial diseases of tomato in the Czech and Slovak Republics and lack of streptomycin resistance among copper-tolerant bacterial strains. *Crop. Prot.* **1995**, *14*, 267–270, doi:10.1016/0261-2194(94)00010-6.
43. Beran, P.; Mraz, I.; Kokoskova, B.; Bohata, A. Monitoring the occurrence of bacterial spot of tomato and pepper in the Czech Republic and development of new PCR primers for detection of *Xanthomonas vesicatoria*. *Eur. J. Plant Pathol.* **2015**, *141*, 617–621, doi:10.1007/s10658-014-0554-y.
44. Obradovic, A.; Mavridis, A.; Rudolph, K.; Arsenijević, M. Bacterial spot of capsicum and tomato in Yugoslavia. *EPPO Bull.* **2000**, *30*, 333–336, doi:10.1111/j.1365-2338.2000.tb00905.x.
45. Kovachevski, I. Trudove na bulgarskoto prirodizpitatelno druzhestvo. *Sci. Works Bulg. Nat. Soc.* **1936**, *17*, 13–24.
46. Karov, S. *Xanthomonas vesicatoria* (Doidge) Dowson of pepper in Bulgaria. *Bulg. Nat. Soc.* **1965**, *14*, 245–250.
47. Mitrev, S.; Pejcinovski, F. Characterization of *Xanthomonas campestris* pv. *vesicatoria*, causal agent of bacterial spot of pepper, cv. Kurtovska kapija. In *Yearbook for Plant Protection; Society for Plant Protection of Republic of Macedonia: Skopje, Macedonia, 1999; Volume X*, pp. 151–163.
48. Mitrev, S. *Phytopathogenic Bacteria of Pepper in Macedonia*; PSI Institute of Southern Crops Strumica: Strumica, Macedonia, 2001.
49. Bogatzevska, N.; Stoimenova, E.; Mitrev, E. Bacterial and virus diseases spread in Bulgaria and Macedonia on field and greenhouse pepper. *Plant Prot.* **2007**, *18*, 17–21.
50. Kizheva, Y.; Vancheva, T.; Hristova, P.; Stoyanova, M.; Bogatzevska, N.; Moncheva, P. Diversity of *Xanthomonas* spp. causal agents of bacterial spot on pepper and tomato plants in Bulgaria. *Biotechnol. Biotechnol. Equip.* **2011**, *25*, 98–104, doi:10.5504/bbeq.2011.0126.
51. Kizheva, Y.; Vancheva-Ebben, T.; Hristova, P.; Bogatzevska, N.; Moncheva, P. First report of *Xanthomonas euvesicatoria* on tomato in Bulgaria. *C. R. Acad. Bulg. Sci.* **2020**, *73*, 140–146, doi:10.7546/CRABS.2020.01.06.
52. Vancheva, T.; Stoyanova, M.; Tatyozova, M.; Bogatzevska, N.; Moncheva, P. Sub-species diversity of *Xanthomonas euvesicatoria* Bulgarian and Macedonian strains from pepper. *Biotechnol. Biotechnol. Equip.* **2014**, *28*, 592–601, doi:10.1080/13102818.2014.947722.
53. Vancheva, T.; Stoyanova, M.; Tasheva-Terzieva, E.; Bogatzevska, N.; Moncheva, P. Molecular methods for diversity assessment among xanthomonads of Bulgarian and Macedonian pepper. *Braz. J. Microbiol.* **2018**, *49*, 246–259, doi:10.1016/j.bjm.2017.08.011.
54. Timilsina, S.; Jibrin, M.O.; Potnis, N.; Minsavage, G.V.; Kebede, M.; Schwartz, A.; Bart, R.; Staskawicz, B.; Boyer, C.; Vallad, G.E.; et al. Multilocus sequence analysis of xanthomonads causing bacterial spot of tomato and pepper plants reveals strains generated by recombination among species and recent global spread of *Xanthomonas gardneri*. *Appl. Environ. Microbiol.* **2015**, *81*, 1520–1529, doi:10.1128/AEM.03000-14.
55. Bui Thi Ngoc, L.; Vernière, C.; Vital, K.; Guérin, F.; Gagnevin, L.; Brisse, S.; Ah-You, N.; Pruvost, O. Fourteen minisatellite markers for population studies of the citrus canker bacterium, *Xanthomonas citri* pv. *citri*. *Mol. Ecol. Res.* **2009**, *9*, 125–127, doi:10.1111/j.1755-0998.2008.02242.x.
56. Pruvost, O.; Vernière, C.; Vital, K.; Guérin, F.; Jouen, E.; Chiroleu, F.; Ah-You, N.; Gagnevin, L. Insertion sequence- and tandem repeat-based genotyping techniques for *Xanthomonas citri* pv. *mangiferaeindicae*. *Phytopathology* **2011**, *101*, 887–893, doi:10.1094/phyto-11-10-0304.
57. Gironde, S.; Manceau, C. 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.* **2012**, *78*, 3266–3279, doi:10.1128/aem.06655-11.

58. Zhao, S.; Poulin, L.; Rodriguez-R, L.M.; Serna, N.F.; Liu, S.-Y.; Wonni, I.; Szurek, B.; Verdier, V.; Leach, J.E.; He, Y.-Q.; et al. Development of a variable number of tandem repeats typing scheme for the bacterial rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. *Phytopathology* **2012**, *102*, 948–956, doi:10.1094/phyto-04-12-0078-r.
59. N’Guessan, C.A.; Brisse, S.; Le Roux-Nio, A.-C.; Poussier, S.; Kone, D.; Wicker, E. 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* **2013**, *92*, 366–374, doi:10.1016/j.mimet.2013.01.012.
60. Cesbron, S.; Pothier, J.; Gironde, S.; Jacques, M.-A.; Manceau, C. Development of multilocus variable-number tandem repeat analysis (MLVA) for *Xanthomonas arboricola* pathovars. *J. Microbiol. Methods* **2014**, *100*, 84–90, doi:10.1016/j.mimet.2014.02.017.
61. Trujillo, C.A.; Arias-Rojas, N.; Poulin, L.; A Medina, C.; Tapiero, A.; Restrepo, S.; Koebnik, R.; Bernal, A.J. Population typing of the causal agent of cassava bacterial blight in the Eastern Plains of Colombia using two types of molecular markers. *BMC Microbiol.* **2014**, *14*, 161, doi:10.1186/1471-2180-14-161.
62. Poulin, L.; Grygiel, P.; Magne, M.; Gagnevin, L.; Rodríguez-R, L.M.; Serna, N.F.; Zhao, S.; El Rafii, M.; Dao, S.; Tekete, C.; et al. 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.* **2015**, *81*, 688–698, doi:10.1128/aem.02768-14.
63. López-Soriano, P.; Boyer, K.; Cesbron, S.; Morente, M.C.; Peñalver, J.; Palacio-Bielsa, A.; Vernière, C.; López, M.M.; Pruvost, O. Multilocus variable number of tandem repeat analysis reveals multiple introductions in Spain of *Xanthomonas arboricola* pv. *pruni*, the causal agent of bacterial spot disease of stone fruits and almond. *PLoS ONE* **2016**, *11*, e0163729, doi:10.1371/journal.pone.0163729.
64. Gétaz, M.; Krijger, M.; Rezzonico, F.; Smits, T.H.M.; Van Der Wolf, J.M.; Pothier, J.F. Genome-based population structure analysis of the strawberry plant pathogen *Xanthomonas fragariae* reveals two distinct groups that evolved independently before its species description. *Microb. Genom.* **2018**, *4*, e000189, doi:10.1099/mgen.0.000189.
65. Rache, L.; Blondin, L.; Flores, C.; Trujillo, C.; Szurek, B.; Restrepo, S.; Koebnik, R.; Bernal, A.; Vernière, C. An optimized microsatellite scheme for assessing populations of *Xanthomonas phaseoli* pv. *manihotis*. *Phytopathology* **2019**, *109*, 859–869, doi:10.1094/phyto-06-18-0210-r.
66. Levinson, G.; Gutman, G.A. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **1987**, *4*, 203–221, doi:10.1093/oxfordjournals.molbev.a040442.
67. Estoup, A.; Jarne, P.; Cornuet, J.-M. Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol. Ecol.* **2002**, *11*, 1591–1604, doi:10.1046/j.1365-294x.2002.01576.x.
68. Valdes, A.M.; Slatkin, M.; Freimer, N.B. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* **1993**, *133*, 737–749, doi:10.1093/genetics/133.3.737.
69. Jarne, P.; Lagoda, P.J. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **1996**, *11*, 424–429, doi:10.1016/0169-5347(96)10049-5.
70. Angers, B.; Estoup, A.; Jarne, P. Microsatellite size homoplasmy, SSCP, and population structure: a case study in the freshwater snail *Bulinus truncatus*. *Mol. Biol. Evol.* **2000**, *17*, 1926–1932, doi:10.1093/oxfordjournals.molbev.a026294.
71. Thieme, F.; Koebnik, R.; Bekel, T.; Berger, C.; Boch, J.; Büttner, D.; Caldana, C.; Gaigalat, L.; Goesmann, A.; Kay, S.; et al. Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria* revealed by the complete genome sequence. *J. Bacteriol.* **2005**, *187*, 7254–7266, doi:10.1128/jb.187.21.7254-7266.2005.
72. Vancheva, T.; Lefeuvre, P.; Bogatzevska, N.; Moncheva, P.; Koebnik, R. Draft genome sequences of two *Xanthomonas euvesicatoria* strains from the balkan peninsula. *Genome Announc.* **2015**, *3*, e01528–14, doi:10.1128/genomea.01528-14.
73. Rademaker, J.L.; Hoste, B.; Louws, F.J.; Kersters, K.; Swings, J.; Vauterin, L.; De Bruijn, F.J. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: *Xanthomonas* as a model system. *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 665–677, doi:10.1099/00207713-50-2-665.
74. Rademaker, J.L.W.; Louws, F.J.; Schultz, M.H.; Rossbach, U.; Vauterin, L.; Swings, J.; De Bruijn, F.J. A comprehensive species to strain taxonomic framework for *Xanthomonas*. *Phytopathology* **2005**, *95*, 1098–1111, doi:10.1094/phyto-95-1098.
75. Jalan, N.; Aritua, V.; Kumar, D.; Yu, F.; Jones, J.B.; Graham, J.H.; Setubal, J.C.; Wang, N. Comparative genomic analysis of *Xanthomonas axonopodis* pv. *citrumelo* F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. *J. Bacteriol.* **2011**, *193*, 6342–6357, doi:10.1128/jb.05777-11.
76. Potnis, N.; Krasileva, K.; Chow, V.; Almeida, N.F.; Patil, P.B.; Ryan, R.P.; Sharlach, M.; Behlau, F.; Dow, J.M.; Momol, M.; et al. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC Genom.* **2011**, *12*, 146, doi:10.1186/1471-2164-12-146.
77. Jacques, M.-A.; Bolot, S.; Charbit, E.; Darrasse, A.; Briand, M.; Arlat, M.; Gagnevin, L.; Koebnik, R.; Noël, L.D.; Portier, P.; et al. High-quality draft genome sequence of *Xanthomonas alfalfae* subsp. *alfalfae* strain CFBP 3836. *Genome Announc.* **2013**, *1*, 01035–13, doi:10.1128/genomea.01035-13.
78. Gagnevin, L.; Bolot, S.; Gordon, J.L.; Pruvost, O.; Vernière, C.; Robene, I.; Arlat, M.; Noël, L.D.; Carrère, S.; Jacques, M.-A.; et al. Draft genome sequence of *Xanthomonas axonopodis* pv. *allii* strain CFBP 6369. *Genome Announc.* **2014**, *2*, e00727–14, doi:10.1128/genomea.00727-14.
79. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580, doi:10.1093/nar/27.2.573.
80. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797, doi:10.1093/nar/gkh340.

81. Nascimento, M.; Sousa, A.; Ramirez, M.; Francisco, A.P.; Carriço, J.A.; Vaz, C. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics* **2017**, *33*, 128–129, doi:10.1093/bioinformatics/btw582.
82. Francisco, A.P.; Vaz, C.; Monteiro, P.T.; Melo-Cristino, J.; Ramirez, M.; Carriço, J.A. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics* **2012**, *13*, 87, doi:10.1186/1471-2105-13-87.
83. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2014**, *2*, e281, doi:10.7717/peerj.281.
84. Kalinowski, S.T. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic diversity. *Mol. Ecol. Notes* **2005**, *5*, 187–189, doi:10.1111/j.1471-8286.2004.00845.x.
85. Excoffier, L.; Laval, G.; Schneider, S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **2005**, *1*, 47–50, doi:10.1177/117693430500100003.
86. Jibrin, M.O.; Timilsina, S.; Potnis, N.; Minsavage, G.V.; Shenge, K.C.; Akpa, A.D.; Alegbejo, M.D.; Beed, F.; Vallad, G.E.; Jones, J.B. First report of atypical *Xanthomonas euvesicatoria* strains causing bacterial spot of tomato in Nigeria. *Plant Dis.* **2015**, *99*, 415, doi:10.1094/pdis-09-14-0952-pdn.
87. Dhakal, U.; Dobhal, S.; Alvarez, A.M.; Arif, M. Phylogenetic analyses of xanthomonads causing bacterial leaf spot of tomato and pepper: *Xanthomonas euvesicatoria* revealed homologous populations despite distant geographical distribution. *Microorganisms* **2019**, *7*, 462, doi:10.3390/microorganisms7100462.
88. Almeida, N.F.; Yan, S.; Cai, R.; Clarke, C.R.; Morris, C.E.; Schaad, N.W.; Schuenzel, E.L.; Lacy, G.H.; Sun, X.; Jones, J.B.; et al. PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. *Phytopathology* **2000**, *100*, 208–215, doi:10.1094/phyto-100-3-0208.
89. Kim, J.-G.; Taylor, K.W.; Mudgett, M.B. Comparative analysis of the XopD type III secretion (T3S) effector family in plant pathogenic bacteria. *Mol. Plant Pathol.* **2011**, *12*, 715–730, doi:10.1111/j.1364-3703.2011.00706.x.
90. Louws, F.; Rademaker, J.; De Bruijn, F. The three Ds of PCR-based genomic analysis of phyto bacteria: diversity, detection, and disease diagnosis. *Annu. Rev. Phytopathol.* **1999**, *37*, 81–125, doi:10.1146/annurev.phyto.37.1.81.
91. Pruvost, O.; Magne, M.; Boyer, K.; LeDuc, A.; Tourterel, C.; Drevet, C.; Ravigné, V.; Gagnevin, L.; Guérin, F.; Chiroleu, F.; et al. 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* **2014**, *9*, e98129, doi:10.1371/journal.pone.0098129.
92. Essakhi, S.; Cesbron, S.; Fischer-Le Saux, M.; Bonneau, S.; Jacques, M.-A.; Manceau, C. Phylogenetic and variable-number tandem-repeat analyses identify nonpathogenic *Xanthomonas arboricola* lineages lacking the canonical type III secretion system. *Appl. Environ. Microbiol.* **2015**, *81*, 5395–5410, doi:10.1128/aem.00835-15.
93. Leduc, A.; Traoré, Y.N.; Boyer, K.; Magne, M.; Grygiel, P.; Juhasz, C.C.; Boyer, C.; Guerin, F.; Wonni, I.; Ouedraogo, L.; et al. Bridgehead invasion of a monomorphic plant pathogenic bacterium: *Xanthomonas citri* pv. *citri*, an emerging citrus pathogen in Mali and Burkina Faso. *Environ. Microbiol.* **2015**, *17*, 4429–4442, doi:10.1111/1462-2920.12876.
94. Ferreira, M.A.S.V.; Bonneau, S.; Briand, M.; Cesbron, S.; Portier, P.; Darrasse, A.; Gama, M.A.S.; Barbosa, M.A.G.; Mariano, R.D.L.R.; Souza, E.B.; et al. *Xanthomonas citri* pv. *viticola* affecting grapevine in Brazil: emergence of a successful monomorphic pathogen. *Front. Plant Sci.* **2019**, *10*, 489, doi:10.3389/fpls.2019.00489.
95. Nakato, G.V.; Rojas, J.L.F.; Vernière, C.; Blondin, L.; Coutinho, T.; Mahuku, G.; Wicker, E. A new Multi Locus Variable Number of Tandem Repeat Analysis Scheme for epidemiological surveillance of *Xanthomonas vasicola* pv. *musacearum*, the plant pathogen causing bacterial wilt on banana and enset. *PLoS ONE* **2019**, *14*, e0215090, doi:10.1371/journal.pone.0215090.
96. Bui Thi Ngoc, L.; Vernière, C.; Jarne, P.; Brisse, S.; Guérin, F.; Boutry, S.; Gagnevin, L.; Pruvost, O.; Guérin, F. From local surveys to global surveillance: three high-throughput genotyping methods for epidemiological monitoring of *Xanthomonas citri* pv. *citri* pathotypes. *Appl. Environ. Microbiol.* **2008**, *75*, 1173–1184, doi:10.1128/aem.02245-08.
97. Schlötterer, C. Evolutionary dynamics of microsatellite DNA. *Chromosoma* **2000**, *109*, 365–371, doi:10.1007/s004120000089.
98. Jibrin, M.O.; Potnis, N.; Timilsina, S.; Minsavage, G.V.; Vallad, G.E.; Roberts, P.D.; Jones, J.B.; Goss, E.M. Genomic inference of recombination-mediated evolution in *Xanthomonas euvesicatoria* and *X. perforans*. *Appl. Environ. Microbiol.* **2018**, *84*, e00136–18, doi:10.1128/aem.00136-18.