agriculture

MDPI

Article

Molecular Evidence for the Presence of Wisteria Vein Mosaic Virus in Italy: Shedding Light on Genetic Diversity and Evolutionary Dynamics of Virus Geographic Populations

Giusy D'Attoma, Angelantonio Minafra, Pasquale Saldarelli and Massimiliano Morelli *

Consiglio Nazionale delle Ricerche, Istituto per la Protezione Sostenibile delle Piante, Sede Secondaria di Bari, 70126 Bari, Italy; giusy.dattoma@ipsp.cnr.it (G.D.); angelantonio.minafra@ipsp.cnr.it (A.M.); pasquale.saldarelli@ipsp.cnr.it (P.S.)

* Correspondence: massimiliano.morelli@ipsp.cnr.it

Abstract: *Wisteria vein mosaic virus* (WVMV) is a member of the genus *Potyvirus*, found to be associated with the Wisteria mosaic disease (WMD). In 2022, severe symptoms resembling WMD were observed on the foliage of a Chinese wisteria (*Wisteria sinensis*) tree growing in Apulia (Italy). The presence of WVMV was ascertained by RT-PCR using the universal degenerate primers NIb2F/NIb3R in the NIb gene. Next, we derived the complete sequence of the CP gene. MAFFT pairwise alignment of the two target regions showed a high amino acid sequence identity of isolate Bari with other WVMV isolates, confirming they belonged to the same species. NIb phylogenetic analysis placed WVMV Bari within the lineage identified in the genus *Potyvirus* as the bean common mosaic virus supergroup. Closer analysis based on the CP gene showed that WVMV Bari was part of a sub-clade gathering all WVMV isolates within a larger clade clustering closely related species. An exploratory analysis of the genetic diversity and possible mechanisms underlying the evolution of geographic populations of WVMV was attempted. The analysis was based on the N-terminal coding region of CP, available for 20 isolates from Europe, Asia, and Oceania. A high diversity, mainly found in the European population, led to the identification of 16 different haplotypes. Based on the dN/dS ratio, the target region appears to be under negative selective pressure. Neutrality tests indicated a potential population expansion in Europe and a recent decrease in Asia. The two populations exhibited a high level of gene flow. We found that WVMV Bari may have originated from a recombination event in the NIb gene. To date, the presence of WVMV had never been reported on *W. sinensis* in Italy, and no molecular information was available on any Italian isolate. Our report draws attention to the further spread of WVMV in the European territory and its rising threat to the ornamental nursery sector.

Keywords: WVMV; selection pressure; population genetics; genetic diversity; gene flow; haplotype diversity; neutrality tests; FastME phylogeny

Citation: D'Attoma, G.; Minafra, A.; Saldarelli, P.; Morelli, M. Molecular Evidence for the Presence of Wisteria Vein Mosaic Virus in Italy: Shedding Light on Genetic Diversity and Evolutionary Dynamics of Virus Geographic Populations. *Agriculture* **2023**, *13*, 1090. https://doi.org/10.3390/agriculture13051090

Academic Editor: Steven Scofield

Received: 21 April 2023
Revised: 15 May 2023
Accepted: 18 May 2023
Published: 19 May 2023

1. Introduction

Wisteria vein mosaic virus (WVMV) is a member of the genus *Potyvirus* in the family *Potyviridae* with flexuous filamentous particles of a modal length of about 760 nm and a positive-sense single-stranded RNA genome of approximately 10 kb [1,2]. WVMV was found to be the causal agent of Wisteria mosaic disease (WMD), which severely affects different *Wisteria* spp. (*W. sinensis*, *W. floribunda*, *W. venusta*, and *W. brachybotrys*) [3,4]. Although WMD was first reported in the United States [5], it was not until 1970 that Bos, in the Netherlands [6], found evidence of the disease etiology and observed particles of a virus. WVMV probably originated in Asia, where it has been reported in association with WMD, in Iran [2,7] and in the Chinese provinces of Beijing [8], Jiangsu [9], and Jangxi [10]. Outside Asia, its presence in plants affected by WMD has been confirmed by molecular methods in Poland [11], the United Kingdom [4], the United States [12], Australia [1], and New Zealand [13]. As in the case of the Dutch isolate [6], earlier reports suggesting the

presence of the virus in Europe, namely in Italy [14] and former Czechoslovakia [15], had only relied on serological data and biological assays.

WVMV is primarily spread through vegetative propagation of infected plant material, but can also be non-persistently transmitted by aphids [16,17], as well as by grafting and mechanical inoculation to a range of indicators [1,18]. The rising worldwide occurrence of WVMV is currently regarded as a significant threat to the wisteria nursery market. Although the virus rarely affects the vigour and longevity of infected plants, chlorosis and mottling induced by WMD considerably impact their foliage, de facto rendering the plants unmarketable [1,9]. Wisteria species, which are woody flowering plants in the family Fabaceae [19], have longstanding popularity in the eastern region of Asia for their ornamental value and their traditional use in oriental homoeopathic medicine [18,20]. The ornamental Chinese wisteria (*W. sinensis*) has gained, in recent years, increasing acclaim among European gardeners and landscape architects for its robust habit and eye-catching and fragrant flowering [9]. In recent years, *W. sinensis* has been identified as the host of several viral agents, including Wisteria badnavirus 1 (WBV1) [21,22], cucumber mosaic virus (CMV) [23,24], and alfalfa mosaic virus (AMV) [25]. However, WVMV is currently regarded as the most serious disease agent adverse to wisteria horticultural production in several countries and has become the subject of a steadily growing body of research.

To date, four complete genome sequences of WVMV isolates [2,3,26] are available in the GenBank database. For a larger number of isolates, the nucleotide sequence of the hypervariable region in the coat protein (CP) gene has been made available for diagnostic purposes. However, these sequences have never been used to analyze the genetic diversity of WVMV variants, which may have important implications for virus populations, host-range, and geographic distribution [27]. Analysing the genetic diversity of WVMV geographic populations and investigating the evolutionary mechanisms that have shaped their population structure may also be important for understanding their emergence [28,29]. Therefore, the analysis of WVMV diversity and its population dynamics are of interest in understanding its epidemiology [29] and preventing its further spread.

This paper provides the first molecular evidence of the presence of WVMV in Italy, along with the first report on its occurrence in *W. sinensis* in this country. Moreover, our study determines several population genetic indices, together with phylogenetic reconstruction and recombination detection, to describe the WVMV population dynamics and evolutionary patterns.

2. Materials and Methods

2.1. Sample Collection

In 2022, severe symptoms resembling those of WMD (i.e., mosaic mottling, interveinal diffuse chlorotic spots, necrotic flecks, and slight distortions) (Figure 1) were observed on the foliage of a Chinese wisteria (*W. sinensis* Sims) 'Sweet' tree growing in an urban garden in Apulia (Italy). Symptomatic compound leaves were randomly collected from different parts of the canopy and stored in plastic bags kept in an ice box before fast processing. A pooled sample of asymptomatic leaves taken from a *W. sinensis* plant grown in a distant garden, showing no symptoms of disease, was also included in the experiment as a negative control.

2.2. Total RNA Extraction, RT-PCR, Cloning and Sequencing

The presence of WVMV, in association with WMD symptoms, was assessed by a reverse transcription polymerase chain reaction (RT-PCR). To this aim, ribonucleic acids (RNA) were extracted from 250 mg of pooled symptomatic leaves, using RNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) and following the procedure described by Morelli et al. [30]. RNAs were also extracted from the asymptomatic leaf control sample. After denaturation for 5 min at 95 °C along with 0.5 µg of random hexamer p(dN)$_6$ primers (Roche LifeScience, Basel, Switzerland), samples were reverse transcribed for 1 h at 42 °C with 150 U of M-MLV reverse transcriptase (Thermo Fisher Scientific, Waltham, MA, USA).

PCR was performed in a 25-µL reaction mixture containing 2.5 µL of cDNA template, 1 U of DreamTaq™ Green DNA polymerase (Thermo Fisher Scientific, USA), 0.5 µL of 10 mM dNTP mix, 2.5 µL of 10X DreamTaq™ Green Buffer (Thermo Fisher Scientific, USA), and 0.5 µL of each 10 mM forward and reverse primer. Degenerate primers NIb2F/NIb2R, designed by Zheng et al. [31] within the nuclear inclusion protein b (NIb) gene of potyviruses, for their universal detection, were used. Primer sequences and cycling conditions were those reported in the original study [31].

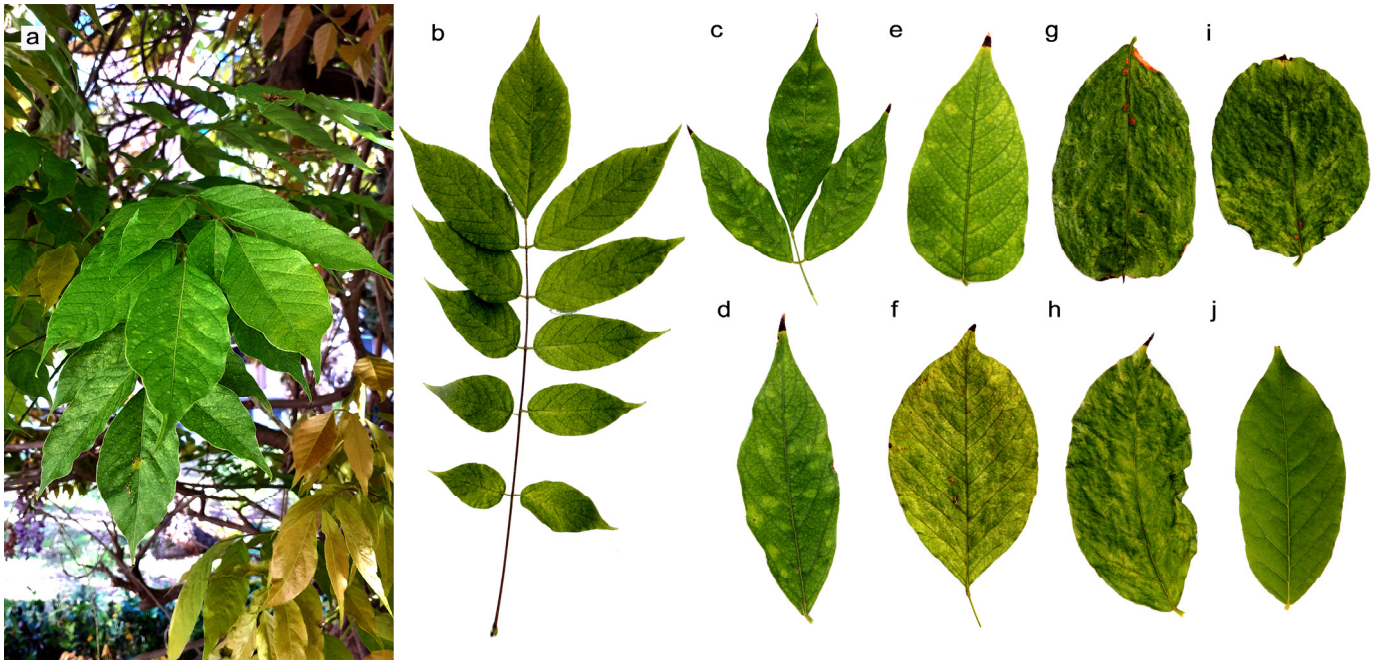

Figure 1. Symptoms of Wisteria mosaic disease (WMD) observed on Chinese wisteria (*Wisteria sinensis* Sims) 'Sweet' leaves, in association with Wisteria vein mosaic virus (WVMV) infection. Milder symptoms on compound leaves varied from yellowing to interveinal diffuse chlorotic spots and irregularly scattered patterns (**a**,**b**). Most severely affected leaflets showed mosaic mottling, ringspots, necrotic flecks, twisting, distortions and tip malformations (**c**–**i**). The leaflet of a healthy plant without symptoms is shown in (**j**) for reference.

A second round of amplification was attempted with a primer pair designed in this study to obtain the complete sequence of the CP gene for use in subsequent phylogenetic and population genetic analyses. The degenerate primers wvcpF (5′-TGTTGTGARTCAGTTTCTCTRC-3′, R = A/G) and wvcpR (5′-ACGACKCGAGATGCTAACTGTG-3′, K = G/T) were designed to span the whole target gene, based on a consensus constructed from CP sequences of WVMV isolates already available in the NCBI Virus database [32]. The PCR reaction mixture was prepared as described below, with the following cycling conditions: denaturation at 95 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, 57 °C for 45 s, and 72 °C for 2 min, and final elongation at 72 °C for 7 min. Amplification products were visualized by electrophoresis on a 1.2% (*w*/*v*) TAE agarose gel and stained with GelRed® (Biotium Inc., Fremont, CA, USA).

Amplicons from the NIb2F/NIb3R (350 bp) and wvcpF/wvcpR (912 bp) PCR reactions were cloned into the pSC-A-amp/kan vector (StrataClone PCR cloning kit, Agilent Technologies Inc., Santa Clara, CA, USA). According to the manufacturer's protocol, Taq-amplified PCR products, which contained 3′-adenosine overhangs, were ligated to the vector arms through A-U base pairing, followed by topoisomerase I-mediated strand ligation. The resulting linear molecules were then transformed into a competent cell line engineered to transiently express Cre recombinase (StrataClone SoloPack, Agilent Technologies Inc., Santa Clara, CA, USA). Recombinant clones were identified by blue-white

screening and plasmid DNA was purified with the Plasmid Extraction Mini Kit (Fisher Molecular Biology, Roma, Italy). Plasmid inserts were custom sequenced (Macrogen Europe, Amsterdam, The Netherlands) at their termini with vector-derived primers M13F/M13R. Sequences resulting from 20 clones, for each of the two PCR reactions, were assembled with CLC Genomics Workbench 3.6.5 (CLC bio, Aarhus, Denmark). The analysis of structural and functional domains conserved within the protein sequences was conducted with the Conserved Domain Search service (CD-Search) web-based tool [33].

2.3. Sequence Comparison and Phylogenetic Analysis

2.3.1. Nuclear Inclusion Protein b (Nib) Gene

The available potyvirus Nib-coding gene sequences were retrieved from the NCBI Virus database [32]. The query was initially limited to sequences classified as RefSeq, opting to retain only one sequence for each viral species included in the genus *Potyvirus*. Thereafter, the corresponding sequences of all available WVMV isolates were also included in the list, along with three viral species not classified as RefSeq but available in GenBank, viz *Passiflora chlorosis virus*, *Passiflora mottle virus*, and *Uraria mosaic virus*. In total, sequences belonging to 147 potyvirus isolates were included in the analysis (Table S1), along with one rymovirus isolate, i.e., ryegrass mosaic virus (RGMV), added as an outgroup.

For each sequence, the conserved region within the NIb gene, bracketed by the NIb2F/NIb3R primer sites (350 bp), was extracted through a BLASTn search. Nucleotide sequences were translated into amino acid sequences using the BioEdit Sequence Editor version 7.0.9 [34]. Needing to compare homologous genes from different species, the analyses were conducted on protein sequences. This allowed for a more stable and robust comparison, as the phylogenetic signatures disappear much more rapidly from DNA sequences than from encoded proteins [35]. In addition, the use of protein sequences may circumvent potential distortions caused by any nucleotide compositional bias in DNA sequences [36]. The translated sequences were submitted to pairwise Multiple Alignment using Fast Fourier Transform (MAFFT) [37] before generating a two-dimensional diversity matrix using Sequence Demarcation Tool (SDT) software, version 1.2 [38]. A phylogenetic analysis was carried out using the FastME/OneClick workflow available at the Next Generation Phylogeny.fr web service [39]. FastME deduces phylogenetic relationships on the basis of distance algorithms. The method is based on the balanced minimum evolution (BalME) criterion [40], which is the very principle of neighbour-joining (NJ). FastME improves on NJ by performing topological shifts with faster algorithms [41]. Briefly, MAFFT alignment was refined with Block Mapping and Gathering with Entropy (BMGE) [42], allowing for a phylogenetic reconstruction inferred by the FastME 2.0 program [41]. Tree refinement was performed using Subtree Pruning and Regrafting (SPR) topological rearrangement, with the BalME method [40] and a decimal precision for branch length set to 6. Bootstrap branch support was computed for over 1000 replicates. The inferred tree was rendered using the Newick display utility [43], prior to final editing with the Interactive Tree of Life (iTOL) software, version 6.5.8 [44].

2.3.2. Coat Protein (CP) Gene

Once the results of the phylogenetic analysis in the NIb region were derived, the CP gene was used as the target for the subsequent analyses. The cleavage site between NIb and capsid proteins was predicted within the sequence available for WVMV Bari, based on the conserved motif shared among potyviruses [45] and using the Expasy PeptideCutter tool [46]. Complete sequences of the CP gene, available for those viral species found to be included in the same clade of WVMV in the NIb analysis, were retrieved from the NCBI Virus database [32]. When multiple sequences were available, four were chosen for each viral species, favouring heterogeneity in geographical origin and host plant. In the case of WVMV, all nine complete sequences available were included. The sequence of the onion yellow dwarf virus (OYDV) isolate Yuhang (NP_871747) was included in the analysis as an

outgroup. A total of 71 complete CP sequences (Table S2) were selected and submitted for analysis, using tools and methods already mentioned above for NIb gene analysis.

2.4. Recombination Analysis

MAFFT alignments of WVMV nucleotide sequences of the 350 bp fragment in the NIb gene and the whole CP gene were examined for the presence of intra-species recombination using the Recombination Detection Program, version 4.101 (RDP4) [47]. The sequences of five and nine WVMV isolates for the NIb and CP genes, respectively, were included in the analysis. A full exploratory recombination scan was performed using all seven detection methods implemented in the RDP4 program (RDP, GENECONV, Chimaera, MaxChi, BootScan, SiScan, 3Seq). A Bonferroni-corrected *p*-value cut-off of 0.01 was considered significant.

2.5. Analyses of Genetic Diversity and Population Dynamics

Studies of population genetics were performed on a set of partial CP sequences from 20 WVMV isolates from different geographic origins retrieved from GenBank. The dataset included seven isolates for which the complete sequence of the CP gene was available, plus 13 isolates for which only a partial sequence was reported (Table S2). Therefore, the analyses were focused on a fragment shared between the 20 WVMV isolates and corresponding to a sequence of 453 nt in the 5′ portion of the CP gene, immediately after the NIb/CP cleavage site, encoding aa residues 1–151.

Briefly, the target full or partial WVMV nt sequences (Table S2) were retrieved from the NCBI Virus database [32], manually edited using the BioEdit Sequence Editor version 7.0.9 [34] and submitted to MAFFT multiple alignment [37] with default settings.

The genetic distinction of WVMV populations was investigated according to their geographic origin, thus grouping isolates from Asia (China, Iran), Europe (Italy, United Kingdom, Poland, The Netherlands), and Oceania (New Zealand, Australia). The software DNA Sequence Polymorphism (DnaSP) version 6.12.03 [48] was used to perform all genetic diversity analyses and estimate population genetic parameters where not otherwise specified. The haplotype diversity (Hd), the number of polymorphic segregation sites (S), the number of singleton variable sites (SIs), the number of parsimony informative sites (PIs), the total number of mutations (η), the average number of pairwise nucleotide differences within population (K), and the average pairwise nucleotide distance between sequences assuming random samples (π) were computed. The π diversity was also calculated based on the sliding window method to assess the stepwise diversity across the 453 nt-CP target region. In this method, a window of known size (100 bp) was moved along the sequences in steps. The number of steps was set to 25 bp. The π parameter was calculated in each window, and the value was assigned to the nucleotide at the midpoint.

The average synonymous (dS) and non-synonymous (dN) substitutions per site and the dS/dN ratio (ω) were estimated by the Jukes and Cantor (JC) correction [49], performed using the simplification indicated in Nei and Miller [50]. Amino acid polymorphism analysis and identification of polymorphic codons were carried out using the Unipro UGENE software, version 44.0 [51].

Genetic differentiation and gene flow between WVMV geographic populations were assessed in pairwise comparisons and expressed in terms of fixation index (F_{ST}) and number of migrants per generation (Nm). The F_{ST} index estimates the amount of genetic variance within a subpopulation with respect to the overall genetic diversity and is based on Wright's *F* statistic [52,53]. Nm was calculated as $[1/F_{ST}) - 1]/4$, according to Slatkin [54].

In addition, other genetic differentiation indices were included in the analysis: the average proportion of nucleotide differences between populations (Kxy), the average number of nucleotide substitutions per site between populations (Dxy), and the number of net nucleotide substitutions per site (Da). Tajima's *D* [55] and Fu-Li's *F* [56] neutrality tests were used to evaluate the hypothesis of selection forces acting on the CP gene and/or evolutionary events shaping WVMV geographic population sizes.

3. Results

3.1. Sequence Comparison and Phylogenetic Analysis

3.1.1. Nuclear Inclusion Protein b (NIb) Gene

RT-PCR amplification was carried out using degenerate primer set NIb2F/NIb2R, originally designed by Zheng et al. [31] within a highly conserved region of the NIb gene of potyviruses. All the amplifications from the symptomatic plant resulted in an amplicon of the expected size (350 bp), while no positive reaction was achieved from the asymptomatic plant. Sequencing of the 20 clones obtained revealed a unique consensus sequence (OP381183.1), predicted to code for an open reading frame (ORF) of 116 aa. The deduced amino acid sequence (UXP87281.1) corresponded to the conserved domain cd23175, which constitutes the catalytic core domain of RNA-dependent RNA polymerase (RdRp) of RNA viruses belonging to the family *Potyviridae*. At positions 42–56, the active site formed by the conserved polymerase motif A (VYCHADGSQFDSSLT), which is critical for the recognition and binding of nucleoside triphosphates and metal ion cofactors [57,58], was identified by CD-Search.

Matrix calculation, based on MAFFT pairwise alignment with isolates of 143 species members of the *Potyvirus* genus, showed that WVMV Bari shared a high level of amino acid (aa) identity (96.55–100%) with the sequences of the corresponding NIb region, available from other WVMV isolates in the NCBI database (Figure 2, Table S3). As for other viruses, the RefSeq of 13 other species showed more than 90% aa identity with WVMV-Bari, bean common mosaic necrosis virus (BCMNV, NP_734357), and Paris mosaic necrosis virus (PMNV, YP_009667127) being the most similar (93.97% and 93.10%, respectively) (Table S4).

Phylogenetic analysis was conducted with NIb sequences retrieved from five isolates of WVMV and 142 other isolates, each representing a distinct potyvirus species. One isolate of RGMV, belonging to the genus *Rymovirus*, was also included and used as an outgroup. The FastME phylogenetic tree allowed 14 distinct lineages to be identified (Clade 1–14, Figure 3). WVMV Bari clustered with the other WVMV isolates in a large clade that included 25 species (Clade 1, Figure 3). In the sub-clade clustering all the WVMV sequences, isolate Bari appeared most closely related to isolate Ir (QLK02401) from Iran, with which it shared 100% aa identity. Interestingly, the cowpea aphid borne mosaic virus (CABMV) isolate Z (NP_734387), representing a species other than WVMV, was also clustered in this sub-clade. The aa sequence identity of CABMV-Z with WVMV isolates ranged from 88.79 to 91.38% (Table S3).

3.1.2. Coat Protein (CP) Gene

RT PCR using primer pair wvcpF/wvcpR, designed for this study, amplified from the symptomatic tree a fragment of 912 bp in a region identified as lying between the terminal end of the NIb gene and the 3′ untranslated region (UTR). Within this fragment (OP748400.3), we could derive the complete sequence (846 bp) of the gene coding for the CP. CD-Search could identify within the 282-aa deduced sequence (WAD86938.3) the conserved domain cl02961 (CP aa 48–277), known as the signature sequence of the potyvirus CP superfamily (pfam00767). The DAG motif was found at two repeated sites in the N-terminal region (CP aa 10–12 and 50–52), which is deemed critical for CP interaction with helper component proteinase (HC-Pro) and potyvirus transmission by aphids [3,59,60]. The NAG triplet, another conserved motif in potyvirus CPs, which is thought to be crucial for aphid-mediated transmission [61,62], was also found near the N-terminal end (CP aa 18–20).

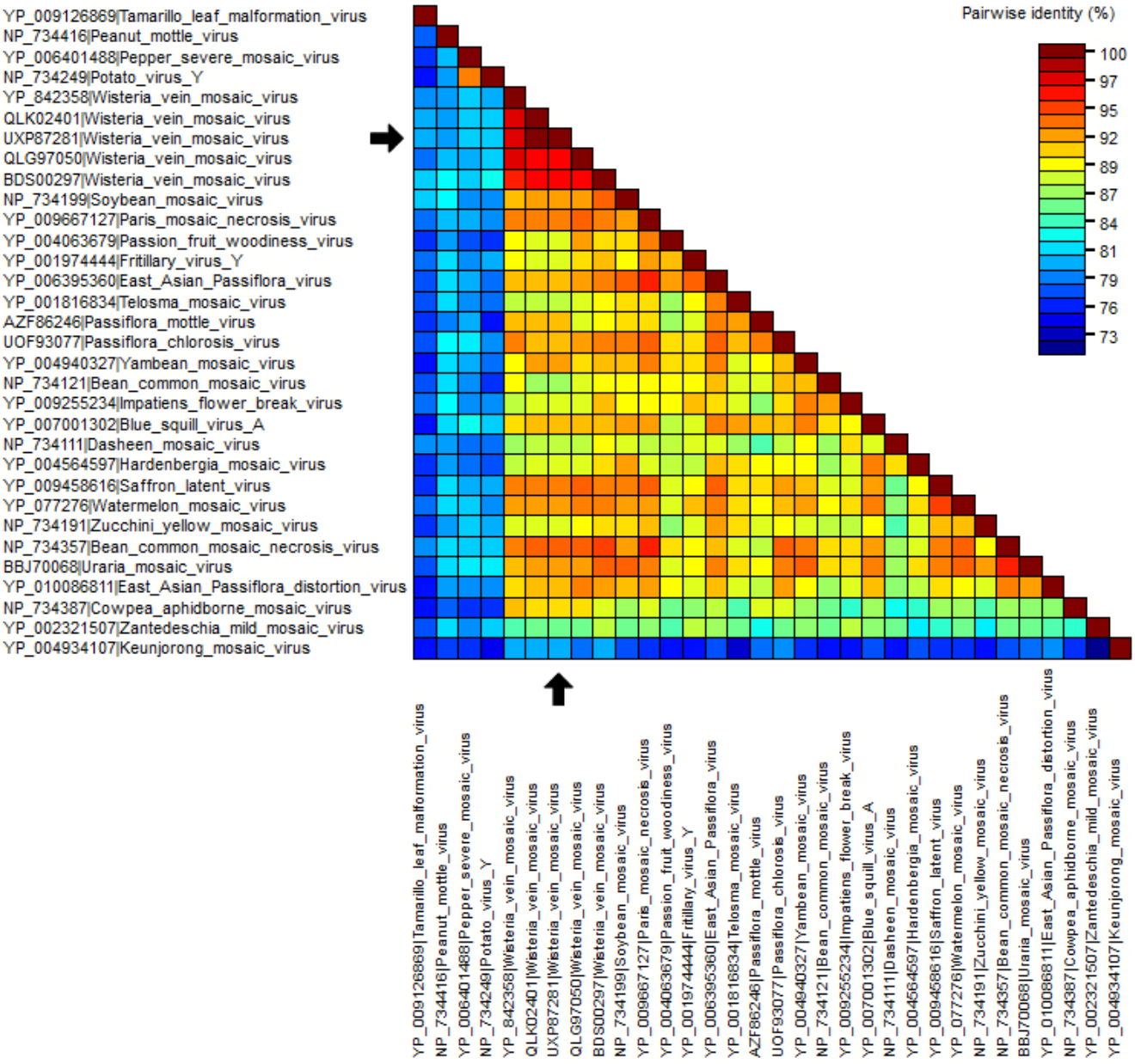

Figure 2. Distance matrix of pairwise MAFFT-based comparison of the partial sequence (116 aa) of the nuclear inclusion protein b (NIb) protein of species members of the genus *Potyvirus*. The analysis was performed using Sequence Demarcation Tool (SDT) software, version 1.2. Only sequences sharing an aa identity ≥80% with the WVMV isolate Bari (UXP87281, indicated by the arrows) are shown. The full matrix can be found in Table S3. All sequences are identified by their accession number and species name. Detailed information on individual isolates can be found in Table S1.

The VSLQS residues found at the NIb/CP junction are consistent with the V[yshr][Hlf]Q/[sa] consensus sequence described by Goh and Hahn [45], comparing 135 potyvirus species. In particular, the site where the proteolytic cleavage occurs was represented by the amino acid residues Gln (Q) and Ser (S). This finding was confirmed by in silico prediction, and is in line with the vast majority of potyvirus sequences [45] and all WVMV isolates described so far.

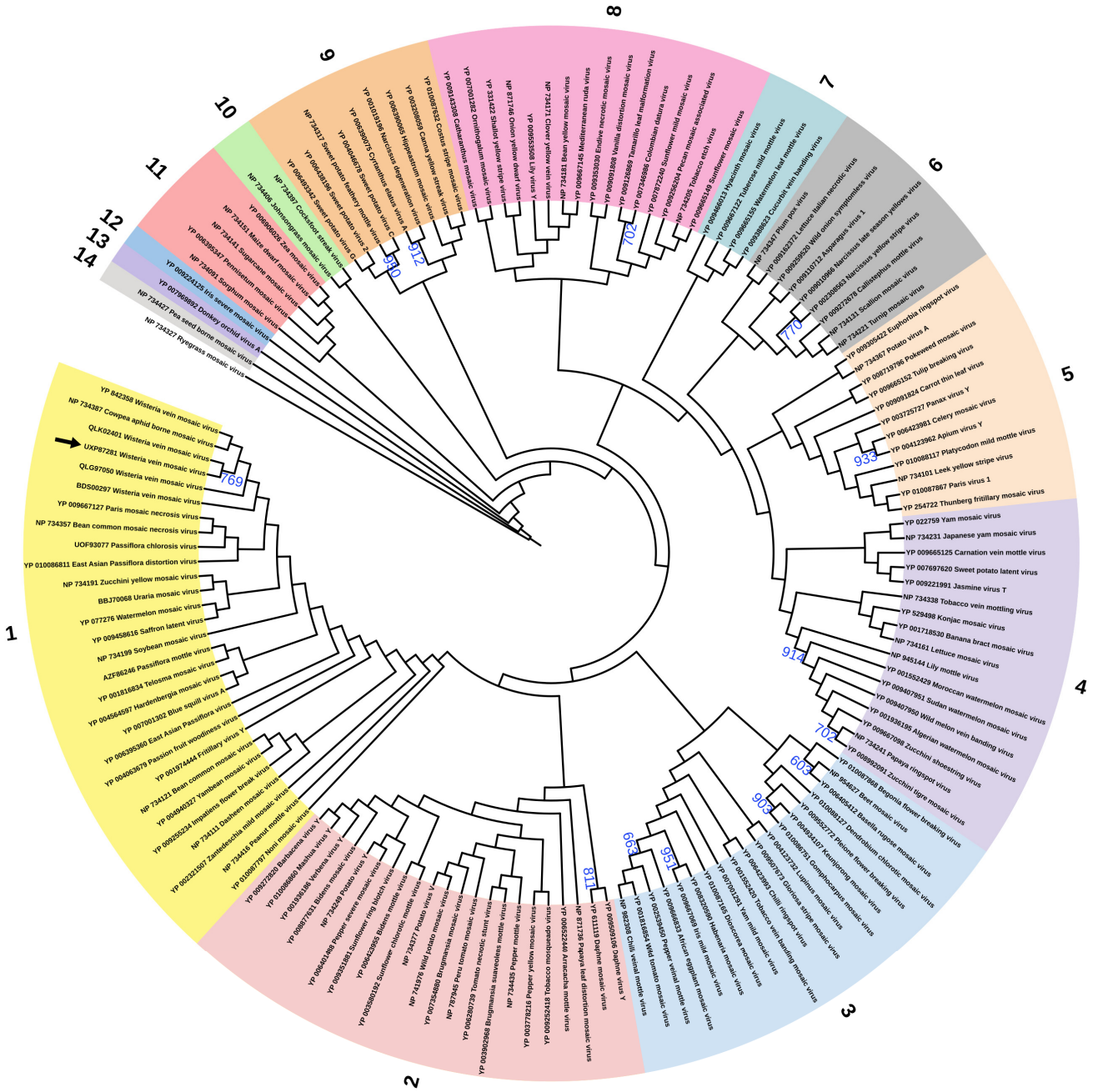

Figure 3. Radial cladogram based on the partial sequence (116 aa) of the nuclear inclusion protein b (NIb) of 147 isolates representative of 143 species members of the genus *Potyvirus*. Phylogenetic reconstruction, based on the balanced minimum evolution (BalME) criterion [40], was performed with FastME/OneClick workflow, available at NGPhylogeny web service, and tree editing was carried out with Interactive Tree of Life (iTOL) software, version 6.5.8. The branch length is unscaled. Bootstrap branch support was computed over 1000 replicates. Bootstrap values higher than 500 are shown above the branches. The tree was rooted using ryegrass mosaic virus (RGMV, genus *Rymovirus*) as an outgroup. A total of 14 clades were found, based on clustering extending from the most distant node. Distinct clades are numbered from 1 to 14 and highlighted in different colours. The arrow indicates the WVMV isolate Bari, the object of this study. All sequences are identified by their accession number and species name. Detailed information on individual isolates can be found in Table S1.

We further compared the CP sequence of WVMV Bari with that of other potyvirus species found to be in the same clade (Clade 1, Figure 3) in the phylogenetic analysis related to the NIb gene. For this purpose, a total of 71 isolates representing 23 different species were included in the analysis. MAFFT pairwise alignment showed that WVMV Bari shared the highest aa sequence identity (100%) with the Iranian isolates Ir (QLK02401) and ThW2259 (AXR98215), followed by the Australian isolate AUS (AAN76998, 99.65%) and three isolates from China, namely Beijing (YP_842359, 96.81%), BJ (AAS89818, 96.81%), and YZ (QBK46786, 96.75%). The closest sequences we identified, apart from WVMV isolates, belonged to soybean mosaic virus (SMV), with aa identities ranging from 81.51% (isolate Mashhad, UQB67141) to 81.89% (isolate CC196, BDC30334), to Uraria mosaic virus (UMV) isolate OC (BBJ70068, 80.73%), and to four calla lily latent virus (CLLV) isolates, with aa identities comprised between 79.79% (isolate 3c27, ABO16409) and 80.50% (isolate BM19, ABO16408) (Figure 4, Tables S5 and S6).

It is worth mentioning that we found two WVMV isolates that shared a CP aa identity far less than 90% with WVMV Bari. This was the case for the South Korean isolates CeJH (BDS00297, 81.56%) and JEBUp (QLG97050, 80.14%) (Table S6). However, these two isolates show some anomalies compared to other known isolates of WVMV, and their classification within the species may be questionable. The JEBUp isolate is, to our knowledge, the only isolate of WVMV ever described in Glycine max (soybean), as it was sequenced, by an NGS approach, from a pooled library of 15 soybean plants, coinfected with different viruses [26]. In our analysis, the CP sequence of JEBUp resulted very similar to CeJH (95.05%) (Figure 4, Table S5). However, its percentage identity with the other WVMV isolates was lower than that with SMV isolates. Interestingly, the aa sequence identity of JEBUp with SMV isolates (81.89–82.64%) (Table S5) is at the threshold of species demarcation within the family Potyviridae, if one relies on the criterion reported by Adams et al. [63]. The whole genome sequence (MT603851.1) is available for the JEBUp isolate. However, when we checked its similarity with the other available full-length sequences, we noticed that JEBUp is 100% identical, at nucleotide (nt) level, to another putative species, isolated by the same authors, and named soybean virus A (isolate JB, MH428831), but never reported elsewhere. The complete genome of JEBUp shows a high nt sequence identity with CeJH (LC729727.1, 87.25%), but the identity with the other WVMV isolates is much lower, around 79%, slightly different from that with other species. All these elements make the attribution to the Wisteria vein mosaic virus species of JEBUp and CeJH controversial. Little information is available on the latter, which was isolated from jack bean (Canavalia ensiformis), but the sequence similarity and the same geographic origin might suggest that CeJH is more closely related to JEBUp, and both are distinct from the WVMV group.

All these findings were corroborated in the subsequent phylogenetic analysis. The phylogenetic tree derived from the CP sequence of the 71 isolates described above allowed for the identification of seven distinct clades (Clades a–g, Figure 5). In particular, WVMV Bari formed part of a sub-clade comprising all WVMV isolates within a larger clade (Clade e, Figure 5), which also included isolates of UMV, CLLV, SMV, and watermelon mosaic virus (WMV). The WVMV sub-clade also comprised the two divergent isolates JEBUp and CeJH, although these formed a distinct sub-group with a higher genetic distance. The WVMV Bari isolate showed a closer relationship to the Iranian isolates Ir and ThW2259, and to a slightly lesser extent, the Australian isolate AUS. The three Chinese isolates, Beijing, BJ, and YZ, formed a distinct, but close cluster.

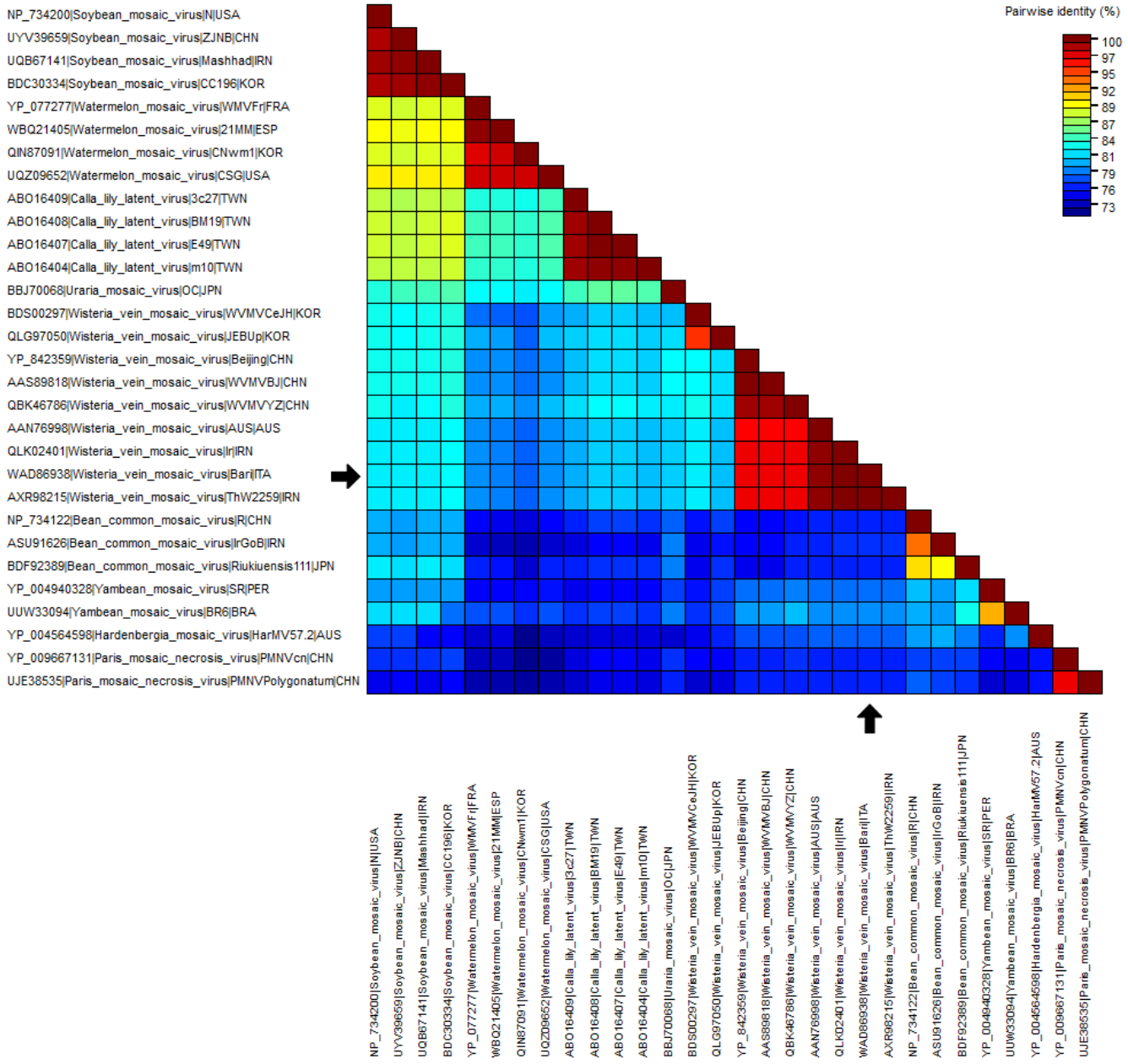

Figure 4. Distance matrix of pairwise MAFFT-based comparison of the complete sequence of the coat protein of species members of the genus *Potyvirus*. The analysis was performed using Sequence Demarcation Tool (SDT) software, version 1.2. Only sequences sharing an aa identity ≥75% with the WVMV isolate Bari (WAD86938, indicated by the arrows) are shown. The full matrix can be found in Table S5. All sequences are identified by their accession number, species and isolate name, and country of origin. Three-letter codes are used for country names, as defined in ISO 3166-1 standard [64]. Detailed information on individual isolates can be found in Table S2.

Figure 5. A rectangular phylogram based on the complete sequence of the coat protein of 71 isolates representative of 23 species members of the genus *Potyvirus*. Phylogenetic reconstruction, based on the balanced minimum evolution (BalME) criterion [40], was performed with FastME/OneClick workflow available at NGPhylogeny web service, and tree editing was carried out with Interactive

Tree of Life (iTOL) software, version 6.5.8. Bootstrap branch support was computed for over 1000 replicates. Bootstrap values above 500 are shown above the branches. The branch length is scaled by evolutionary distance (substitution/site). The tree was rooted using onion yellow dwarf virus (OYDV) isolate as an outgroup. A total of seven clades were found, based on clustering extending from the most distant node. Distinct clades are lettered from a to g. The WVMV isolate Bari, the object of this study, is highlighted in green. The scale bar represents the branch length values. All sequences are identified by their accession number, species and isolate names, and country of origin. Three-letter codes are used for country names, as defined in ISO 3166-1 standard [64]. Detailed information on individual isolates can be found in Table S2.

3.2. Analyses of Genetic Diversity and Population Dynamics

3.2.1. Genetic Diversity and Amino Acid Changes

An analysis of the genetic diversity within the N-terminal coding region of the CP gene of WVMV (aa residues 1–151) was conducted on 20 isolates originating from three different geographic macro-regions (Asia, n = 5; Europe, n = 13; Oceania, n = 2). The decision was made not to include the CeJH and JEBUp isolates in the analysis, as a preliminary MAFFT alignment had shown for both an nt sequence identity far below 80% (68.67–72%) with the other isolates (Table S5). Analysis of the 453 bp long sequence revealed that in these 20 isolates, 360 sites were invariable or monomorphic. The remaining 93 sites, yielding 102 mutations in total, were polymorphic. The occurrence of trimorphic (frequency 7.53%) or tetramorphic (1.08%) variants was rare, the nucleotide substitutions being bi-morphic in the majority of cases (91.4%). A variable site was identified as a singleton if it contained at least two types of nucleotides, with, at most, one occurring multiple times. If a variable site contained at least two nucleotides with a minimum frequency of two, it was considered parsimony informative. In our analysis, 25 polymorphic sites were singleton variable and 68 were parsimony informative.

This genetic differentiation led to the identification of 16 different haplotypes (Hap_1 to Hap_16). The most represented haplotypes were Hap_2 (n = 3: isolates Sar 10, Sar 8, JW 2014), Hap_10 (n = 2: isolates Wis 2B, AUS), and Hap_12 (n = 2: isolates BJ, Beijing) (Table S7). The remaining 13 haplotypes were each found in a single isolate only. With reference to their geographic origin, we found that almost all haplotypes were region-specific, apart from the aforementioned Hap_2, identified in two Italian (Sar 10, Sar 8) and one Chinese (JW_2014) isolates, and Hap_10, which was also sequenced from two different continents, common to Dutch (Wis 2B) and Australian (AUS) WVMV isolates. The European isolates had a higher number of polymorphic sites and mutations (S = 92, η = 100) than the Asian (S = 65, η = 67) and Oceanic (S = 55, η = 55) isolates. Interestingly, all 55 polymorphic sites in the Oceania group were singleton variable, while among the Asian isolates, a large prevalence of parsimony informative sites was found (57 out of 65 total segregating sites).

The haplotype diversity (Hd), which represents the probability that two randomly sampled alleles are different [65], among all 20 isolates was very high and close to 1 (0.974), indicating high genetic diversity within the WVMV population. Concerning nucleotide diversity (π), i.e., the average number of nucleotide differences per site in pairwise comparisons [65], an overall value of 5.32% was observed. The average number of nucleotide differences between sequences (K) was 24.116. When referring to geographic population groups, we found that the European and Oceanic isolates had higher Hd values (1.0) than those found in Asia (0.9). In turn, nucleotide diversity and difference values were lower in Europe (π = 6.19%, K = 28.051) than in Asia (π = 8.37%, K = 37.9), and even lower compared to Oceania (π = 12.14%, K = 55.0) (Table 1).

Table 1. Genetic diversity measures of Wisteria vein mosaic virus (WVMV) isolates, grouped by geographic origin, and based on a 453 nt fragment of their coat protein gene, encoding for amino acid residues 1–151.

Geographic Group	n	H	Hd	S	SVs	PIs	Pcd	η	K	π
Europe	13	13	1.000	92	27	65	9	100	28.051	0.062
Asia	5	4	0.900	65	8	57	6	67	37.900	0.084
Oceania	2	2	1.000	55	55	0	1	55	55.000	0.121
Total	20	16	0.974	93	25	68	13	102	24.116	0.053

n: number of isolates; H: number of haplotypes; Hd: haplotype diversity; S: number of segregating/variable sites; SVs: number of singleton variable sites; PIs: number of parsimony informative sites; Pcd: number of polymorphic codons; η: total number of mutations; K: average number of nucleotide differences between sequences; π: average pairwise nucleotide diversity (per site). All genetic measures and statistical tests were computed on DNA Sequence Polymorphism (DnaSP) software version 6.12.03.

Sliding window analysis, using a step size of 25 bp, showed that in all geographic populations, the highest nucleotide diversity (π) was found in the range between nt positions 325 and 375 (Figure 6). It must be remarked that the measure of diversity observed in Oceania should be weighed against the small number of isolates available (n = 2), even though the two haplotypes presented interesting traits, as also shown by the aforementioned high number of singleton variable sites.

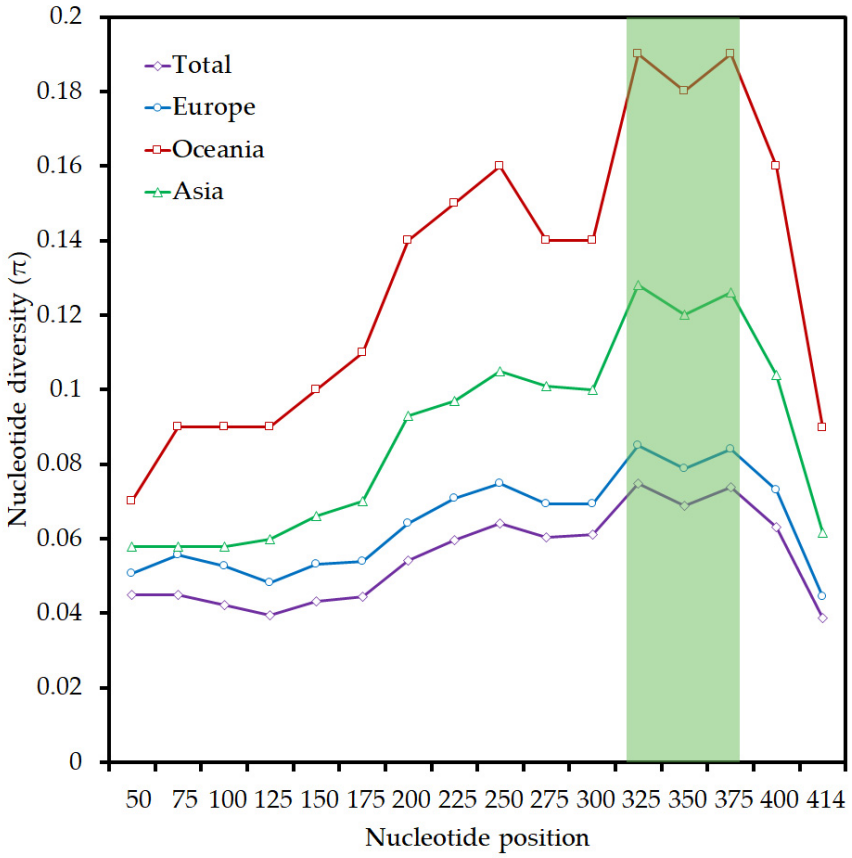

Figure 6. Sliding window plot of the nucleotide diversity per site (π) comparing the level of genetic diversity among the sequences of a 453 nt fragment of the coat protein gene of Wisteria vein mosaic virus (WVMV) isolates. Sequences were grouped by geographic origin (Asia, n = 5; Europe, n = 13; Oceania, n = 2; total, n = 20). The π values were computed on DNA Sequence Polymorphism (DnaSP) software version 6.12.03, with a 100 bp window length and a 25 bp step size. The graphic plot was created with Microsoft Excel® 2016 spreadsheet. The nucleotide range showing the highest π values is highlighted in green.

Protein sequence analysis revealed 13 polymorphic residue positions (Figure S1). All these 13 amino acid changes were dimorphic and occurred at positions 7 (Q/H), 15 (S/Q), 16 (K/R), 18 (N/T), 19 (A/T), 29 (V/M), 72 (V/L), 86 (D/E), 110 (A/V), 113 (G/S), 118 (S/D), 123 (E/S), and 124 (I/V). Although none of the 13 codon polymorphisms were common to all three geographic groups, five amino acid changes were shared between European and Asian WVMV isolates (V29M, V72L, D86E, S118D, E123S, I124V). The amino acid mutations Q7H and A110V were found only in Oceania and Asia, respectively, while five mutations (S15Q, K16R, N18T, A19T, G113S) were exclusive to European isolates. Interestingly, the Italian isolate Sar 5 harboured eight of the thirteen mutant codons identified overall, and four mutations (S15Q, K16R, N18T, A18T) were found only in its sequence. European isolates showed a higher number of mutant codons (9) when compared to Asian (6) and Oceanic isolates (1) (Figure S1).

3.2.2. Analysis of Selection Pressure and Tests of Neutrality

When we sought evidence of selection pressure, it was found that the average ratio between non-synonymous and synonymous substitutions (dN/dS = ω) was 0.051 for all WVMV sequences, suggesting that the target region of the WVMV CP gene is under a strong negative evolutionary constraint. The ω rate was almost the same and equal among all geographic groups, ranging between 0.042 (Asia) and 0.049 (Europe) when analysed at the intra-population level. The existence of purifying selection that could limit variability in WVMV CP was also suggested by the fact that all three geographic groups had lower values of non-synonymous than synonymous substitutions.

Tajima's *D* test was applied to estimate the hypothesis of neutral selection acting on the cp gene. The value of Tajima's *D* statistic was negative (−0.67) and non-significant ($p > 0.10$) when the overall population was considered (Table 2). A negative, not significant value (−0.60, $p > 0.10$) was also observed for the European group. The Asian isolate group, conversely, showed a positive Tajima's *D* value (1.416); however, it was not significant ($p > 0.10$). A low population differentiation among European isolates was also supported by the negative value resulting from Fu- Li's *F* test (−0.186). In contrast, the Asian group recorded a positive Fu-Li's *F* value (1.528). However, both groups showed no statistically significant departure from neutrality ($p > 0.10$). Tajima's *D* and Fu-Li's *F* statistical tests could not be performed for Oceanic isolates due to the limited number of available isolates, as at least four sequences were needed.

Table 2. Evaluation of selection pressure and tests of neutrality of Wisteria vein mosaic virus (WVMV) isolates, grouped by geographic origin and based on a 453 nt fragment of their coat protein gene, encoding for amino acid residues 1–151.

Geographic Group	n	dN	dS	ω	Tajima's *D*	Fu-Li's *F*
Europe	13	0.014	0.295	0.049	−0.598 ^{ns}	−0.186 ^{ns}
Asia	5	0.019	0.442	0.042	1.417 ^{ns}	1.528 ^{ns}
Oceania	2	0.032	0.728	0.044	n.a.	n.a.
Total	20	0.012	0.245	−0.051	−0.674 ^{ns}	−0.413 ^{ns}

n: number of isolates; dN: average number of non-synonymous mutations per non-synonymous site; dS: average number of synonymous mutations per synonymous site; ω: dN/dS, average ratio between non-synonymous and synonymous mutations in pairwise sequences; ns: not significant (*p*-value > 0.10); n.a.: not available, due to limited number of isolates available. All genetic measures and statistical tests were computed on DNA Sequence Polymorphism (DnaSP) software version 6.12.03.

3.2.3. Genetic Differentiation and Gene Flow

The measure of genetic differentiation and gene flow were computed through fixation indices (F_{ST}) and the number of migrants per generation (Nm), respectively. The F_{ST} values between Europe and Oceania and between the latter population and the Asian population were negative, indicating the absence of differentiation at these loci [23]. The comparison between the European and Asian populations showed the existence of gene differentiation, albeit modest ($F_{ST} = 0.126$), with a very high gene flow (Nm = 1.732).

The highest degree of genetic differentiation between European and Asian isolates was also confirmed by the highest values of other inter-population genetic differentiation parameters. The average number of pairwise nucleotide differences (Kxy = 37.615), nucleotide substitution per site (Dxy = 0.083), and net nucleotide substitution per site (Da = 0.010) were higher between European and Asian populations than in the pairwise comparisons Europe/Oceania and Asia/Oceania (Table 3).

Table 3. Indices of genetic differentiation and gene flow between geographic populations of Wisteria vein mosaic virus (WVMV) isolates based on a 453 nt fragment of their coat protein gene, encoding for amino acid residues 1–151.

Population A	Population B	F_{ST}	Nm	Kxy	Dxy	Da
Europe (n = 13)	Asia (n = 5)	0.126	1.732	37.615	0.083	0.010
Europe (n = 13)	Oceania (n = 2)	−0.237	−1.306	33.577	0.074	−0.018
Asia (n = 5)	Oceania (n = 2)	−0.498	−0.752	31.000	0.068	−0.034

n: number of isolates; F_{ST}: Wright's *F* fixation index; Nm: number of migrants per generation, $[(1/F_{ST}) - 1]/4$; Kxy: average proportion of nucleotide differences between populations; Dxy: average number of nucleotide substitutions per site between populations; Da: number of net nucleotide substitutions per site between populations. All genetic measures were computed on DNA Sequence Polymorphism (DnaSP) software version 6.12.03.

3.3. Recombination Analysis

Recombination analysis carried out using the algorithms implemented in the RDP4 program identified potential recombination events between WVMV isolates, which occurred in both regions target of our investigation (Table 4). Interestingly, one putative recombination event was found in the isolate object of this study (Bari, OP381183) within the NIb gene (nt positions 34–321) and confirmed by MaxChi (*p*-value 8.39×10^{-3}) and 3Seq (*p*-value 3.44×10^{-6}) algorithms. The Iranian isolate Ir (MN514947) was identified as the putative minor parental sequence; however, the major parent was labelled as "unknown". A putative recombination event was identified also in the CP gene (nt positions 492–540) and involved isolate JEBUp (MT603851), recognized as a possible recombinant between CeJH (LC729727) and Beijing (NC_007216), which were identified as major and minor parents, respectively. In this case, the recombination event was detected by the algorithms RDP (*p*-value 1.07×10^{-2}), GENECONV (*p*-value 3.86×10^{-2}), and MaxChi (*p*-value 3.38×10^{-2}).

Table 4. Putative recombination events identified by the detection algorithms of the RDP4 software, version 4.101, within the genes for the NIb and capsid proteins of Wisteria vein mosaic (WVMV) isolates.

Gene	Breakpoint Position (nt)		Recombinant Isolate (Accession)	Parental Isolates (Accession)	
	Beginning	End		Major	Minor
NIb	34	321	Bari (OP381183)	Unknown	Ir (MN514947)
CP	492	540	JEBUp (MT603851)	CeJH (LC729727)	Beijing (NC_007216)

Gene	Detection Algorithm (*p*-Value)						
	RDP	GENECONV	Chimaera	MaxChi	BootScan	SiScan	3Seq
NIb	ns	ns	ns	8.39×10^{-3}	ns	ns	3.44×10^{-6}
CP	1.07×10^{-2}	3.86×10^{-2}	ns	3.38×10^{-2}	ns	ns	ns

ns: not significant (*p*-value > 0.01).

4. Discussion

The presence of a viral agent belonging to the "potato virus Y group", according to the authors' original definition, had been reported in Italy in the 1960s in another wisteria species (*W. floribunda*) [14], which, unlike *W. sinensis*, is native to Japan. At the time, the provisional name Wisteria vein mosaic virus could only be attributed to the authors based on certain similarities observed with another viral isolate concomitantly studied in

Holland [6]. The two isolates could be responsible for a similar aetiology on *Wisteria* spp. infected plants and, to some degree, on herbaceous indicators after mechanical transmission, and were closely related in serological slide tests [14]. As Bos had tentatively named the isolate in his collection 'Wisteria vein mosaic virus', Conti and Lovisolo adopted the same name for the Italian isolate [14]. The evidence, again based only on 'rather incomplete' serological assays [6], electron microscopy observations, and cross-protection trials, was that both isolates could be a viral entity distinct from the closest taxonomic species known at the time, i.e., *Bean common mosaic virus*. However, as the authors themselves admitted [6], the data at their disposal were still not sufficient to define the existence of a new viral species. Since that time and before the current report, to our knowledge, the presence of WVMV in Italy had never been confirmed, and the present study is the first to provide molecular information on an Italian isolate of WVMV. It is worth noting that, in the course of the preparation of this manuscript, and after the deposit of our sequences, other authors submitted to GenBank the sequences of eight isolates of WVMV (OM417219 to OM417226), which were allegedly identified in Italy on *W. sinensis*, in a geographic area 800 km away from our finding. The available sequences only refer to a partial fragment of 668 nt, between the 3′ end of the NIb gene and the 5′ end of the CP gene. Relative to the 488 nt fragment matching with the CP sequence available from WVMV Bari, the nt identity was in the range 87.70–99.39%, rising to 94.44–100% at the aa level. Unfortunately, no further information on these isolates is available, but the detection of WVMV in another Italian region could provide further confirmation of the virus spread in our country. Moreover, these isolates would have been identified in Liguria, a region bordering Piedmont, where Conti and Lovisolo reported their isolate in 1969 [14], which could suggest that WVMV may have been circulating in Italy for more than 50 years.

MAFFT pairwise alignment of WVMV Bari with sequences available from other isolates showed a high level of aa identity in the core region of the NIb protein and the whole capsid protein. This confirms that the Bari isolate should be considered a new strain of the *Wisteria vein mosaic virus* species. We found 13 RefSeqs belonging to other species sharing more than 90% aa identity with WVMV Bari. This was not unexpected, as the NIb gene has been described as one of the most conserved sites in the potyvirus genome [66]. Its core zone, flanked by the NIb2F/NIb3R primers, which we used for detection, is, indeed, one of the most conserved [31]. These species included *Watermelon mosaic virus*, *Soybean mosaic virus*, and *Bean common mosaic virus*, whose close relationship with *Wisteria vein mosaic virus* had already been reported for other isolates [1–3], besides being proven by serological assays [17,67]. Comparison of the CP sequences confirmed the high degree of identity with SMV isolates, as already known, and with the SMV-related UMV strain OC isolated from *Passiflora edulis* [68] and CLLV strains isolated from calla lily (*Zantedeschia* spp.) [69].

Our phylogenetic analysis based on NIb sequences was extended to all potyvirus species included in the NCBI virus database at the time of writing. Adopting the distance-based algorithm FastME [41], which improves over the NJ method [70], we could identify 14 distinct clades, which confirm most, but not all, of the lineages resulting from genome-scale maximum likelihood (ML) analyses proposed in earlier studies [71,72]. According to Herath et al. [73], clades were recognized based on the group of branches originating from the most distant node, thus suggesting a common lineage ancestor. WVMV isolates were found to be closely related to each other and formed a cluster, supported by high bootstrap values and belonging to a large clade including 25 species. This broad lineage, which we named Clade 1 (Figure 3), is broadly similar to Lineage 3 proposed by Gibbs et al. [72]. This lineage is usually referenced as the bean common mosaic virus (BCMV) supergroup [71], which is estimated to have first emerged more than 3600 years ago in South and East Asia [74–76].

We looked more closely into the phylogenetic relationships within this clade, based on the CP sequence, and extended the characterization to a larger number of isolates. Our findings led to the identification of seven clades within the supergroup, one of which

(Clade e, Figure 5) encompassed all isolates of WVMV in a sub-cluster, along with other sub-clusters gathering isolates of the closely related SMV, WMV, CLLV, and UMV.

The tree we obtained placed the CP sequence of WVMV Bari in close relationship with the Iranian isolates Ir and ThW2259 and the Australian isolate AUS. As noted by Valouzi et al. [16] and Al-Jaberi et al. [2], this "world" group remained neighbouring but separate from that formed by the Chinese geographic subpopulation (isolates YZ, Beijing, and BJ). A more pronounced distance separated a further subgroup, which clustered the two most divergent isolates, CeJH and JEBUp, from South Korea. Our results seem to reinforce the hypothesis that geographic origin may contribute to the evolution of this virus [16], as in other potyviruses [77], although clustering of WVMV based on phylogeography had recently been questioned [2].

As many other factors may contribute to the evolutionary diversification of potyviruses [78,79], the role of host adaptation cannot be underestimated [16]. It has been proposed [2] that the high genetic diversity observed in JEBUp could also be explained by ecological adaptation to a divergent host, the soybean. Over time, the virus would accumulate the necessary variation to adapt to the new host species. This argument could apply to the other divergent isolate, CeJH, also isolated from an unusual host, the jack bean. Conversely, the WVMV isolates showing more conserved sequences and forming tight phylogenetic clusters were all isolated from the same genus, *Wisteria*. They are likely to have undergone similar host adaptive selection [2].

Genetic variations and possible mechanisms lying behind the evolution of geographic populations of WVMV were explored by analysing the 5′ terminal variable region (aa 1–151) of the CP gene of 20 isolates available in the NCBI Virus database [32]. In our polymorphism analysis, we found that nucleotide substitutions affected about 20% of the sites considered. However, over 85% of the nucleotide substitutions observed were silent, leading to synonymous changes in the amino acid sequence. The rare occurrence of non-synonymous variants is in line with what has already been reported on larger-scale analyses of potyvirus sequences [80]. Synonymous mutations are much less likely to be deleterious to protein structure and function than non-synonymous mutations [81,82]. Purifying selection may have acted during the evolution of potyviruses to eliminate non-synonymous mutations to a much greater extent than synonymous mutations [80]. The situation observed for WVMV fits well with an evolutionary model in which purifying selection at the expense of non-silent variants is underway within the population.

The non-synonymous mutations we identified led to the occurrence of 13 polymorphic amino acid residues. Interestingly, two of these changes, A19T and G113S, although representing low-frequency variants (5% and 10%, respectively), constituted substitutions of non-polar amino acids, alanine and glycine, into polar amino acids, threonine and serine. This modification could interfere with the physicochemical traits of the capsid protein, as it involves a switch between hydrophobic and hydrophilic behaviour [83]. However, the biological effect of such mutations should be verified through targeted studies. The occurrence of mutated codons was more frequent in the case of the European WVMV isolates. However, this was mainly due to the high number of mutations found in Sar 5, one of the isolates identified in Liguria, which comprised almost all of the polymorphic codons found in Europe (8 out of 9).

The analysis of the genetic diversity of the 20 isolates we considered led to the identification of a surprisingly high number of different haplotypes. However, only one haplotype, named Hap_2, found in three isolates, and Hap_10 and Hap_12 found in two isolates can be considered, based on their frequency, as major haplotypes. All the others should be regarded as minor haplotypes, as they were limited to a single occurrence. Apart from Hap_12, which could be considered a Chinese haplotype, having been found in the BJ and Beijing isolates, we found no other haplotypes specifically associated with a geographic location. However, this is not unexpected, as the geographic origin is only one of the factors that may help shape the genetic structure of plant viral populations, along with the host plant, mode of transmission, and epidemic history [84]. It is also true that the geographic

branching of WVMV haplotypes may not be complete, as the number of sequences obtained from different geographic origins is still limited. More variants are required to support any further speculation.

The occurrence of a large number of different haplotypes is reflected in the very high values of haplotype diversity (Hd), which confirms a high level of genetic diversity, even more evident within European populations. One might argue that such wide genetic diversity is not equally supported by the nucleotide diversity values (π), which stay in the 0.062–0.121 range, depending on the geographic groups considered. The average value of π (0.053) is lower than the average value measured in the vast majority of potyviruses in genome-scale comparisons [79] or limited to the CP gene, in other species such as *Narcissus late season yellows virus* and *Narcissus yellow stripe virus* [85]. It should be considered, however, that an average pairwise nucleotide diversity higher than 5% remains greater than that reported for viruses used as a standard reference of genetic stability, such as tobacco mosaic virus (TMV) and wheat streak mosaic virus (WSMV) [79]. The situation we observed, with a large number of haplotypes, actually characterized by very minor nucleotide variations, could be related to a population that has recently experienced an expansion phase [28]. Interestingly, this behaviour has been described for viral populations that spread through infected propagating plant material, as was observed in the case of grapevine fanleaf virus (GFLV) [86]. The fact that we observed a lower π nucleotide diversity in European isolates contributes to the hypothesis that WVMV may have a non-European, possibly Asian, origin. Indeed, it is known that pathogens typically exhibit higher genetic diversity in their places of origin [85,87]. In general, it should also be taken into account that the intensity of genetic variation is strongly influenced by sample size. Populations with large sample sizes show wider genetic diversity [83,88]. For this reason, the limited number of available sequences may have skewed the indices of nucleotide diversity we observed.

Sliding window analysis identified an area of higher nucleotide diversity located in a 50-nt portion closer to the N-terminal end of the CP. The existence of hypervariable areas located in the N-terminus of the CP is a feature already described in potyviruses [79]. Based on mutation studies, mainly conducted in plum pox virus (PPV) [89,90], as well as in potato virus Y (PVY) [91], it has recently been proposed that these areas of high genetic flexibility in the N-terminal end of the CP may act as determinants of host adaptation [79].

The selective pressure that could act on the target CP region was investigated in the different geographic groups. The analysis was based on the relative ratio ω, calculated between non-synonymous (dN) and synonymous (dS) substitutions. The values of ω, always found to be less than 1, indicate that the region is placed under negative purifying selection [28]. Nigam et al. [79] proved that most coding regions in the potyvirus genome are under strong negative selection, measuring that negative selection sites were also 10 times more abundant than those placed under positive selection. However, protein-coding regions, which, like CP, are involved in virion particle formation and vector transmission, and which require interactions with cellular machinery, usually contain hypervariable regions, as we have observed, and may display sites of positive selection [79,92].

Tajima's *D* and Fu-Li's *F* neutrality tests applied to determine the growth of geographic populations of WVMV and estimate the neutrality of nucleotide variability in the target CP region showed different results for the European and Asian groups. Negative values were found for both tests in European isolates. This may indicate an excess of low-frequency polymorphisms, potentially caused by a population expansion [83]. In converse, the positive values found in the Asian group could indicate a recent decrease in population size or a balancing selection and a bottleneck, currently maintaining the population at intermediate frequency variants [28,93]. The contraction in the Asian population, reflected by negative Tajima's *D* values, may also have resulted from a recent merger of two distinct populations [94]. Nevertheless, these remain exploratory hypotheses, as we could not find significant *p*-values in any of the groups, and all our inferences remain statistically not

supported. This could be a consequence of the low number of different isolates in each geographic group [85].

When we explored genetic differentiation and the extent of gene flow between different WVMV geographic populations, we found that the parameters of inter-population genetic differentiation showed evidence of a population structuring between the European and Asian groups, not observed in the other combinations. Moreover, these two populations appear to possess a high level of gene flow, as shown by the number of migrants per generation (Nm) far exceeding 1 [48,95]. Physical and geographic distances make long-range virus migration vectored by aphids unlikely. The most plausible hypothesis is that this flow was driven by the circulation of propagation material, probably in the direction from the places of origin of wisteria cultivation to European nursery settlements.

It was probed whether the genetic differentiation of WVMV isolates could also be partially due to the occurrence of recombination events. This would not have been surprising given that natural intraspecific recombination in plant RNA viruses, specifically in the family *Potyviridae* [77], is a very frequent adaptation process. This mechanism compensates for the occurrence of deleterious mutations, and on the other hand, functions as a flywheel for genetic diversity [85,96,97]. Our analysis revealed two single putative recombination events. The first recombination breakpoint was identified in the NIb gene and would have originated the isolate Bari, the object of this study, and the second in the CP gene from which the South Korean JEBUp isolate would have stemmed.

Significantly, both recombination events we identified, if confirmed, could have geographic rationales. In one case, WVMV Bari would have resulted from the recombination event between the Iranian isolate Ir and an unknown minor parent. In the second case, a recombination event restricted to the Far East region would have occurred between the Beijing and CeJH isolates, originating the JEBUp isolate. Similarly to what has been observed for other potyvirus species [98], this type of recombination could have occurred in specific geographic regions, only to remain confined there, as in the case of the South Korean isolate, or to be moved to distant areas by trade in plant material, as in the case of the Italian isolate. The possible involvement of highly divergent isolates, such as CeJH and JEBUp, when proved, would leave the way open for interesting insights. It could be reasoned that, presumably in Eastern Asia, and most likely in herbaceous Fabaceae hosts, WVMV might have undergone recombination events giving rise to isolates, or even to a different species, with genetic and biological characteristics that are on the edge with those of other species, such as SMV. To accurately recognize when and where recombinations occurred would require a much larger number of available sequences. This would allow the identification of recombination hotspots with the support of robust statistical significance [98]. Based on our data, we can only conclude that recombination does not seem very frequent in WVMV, and this does not come unexpectedly. Low recombination frequency is often associated in potyviruses with a narrow host range [85], which, in the case of WVMV, apart from single and questioned detection episodes in *C. ensiformis* and *G. max* [26], seems restricted to the genus *Wisteria*.

The fact that WVMV has so far been reported as naturally occurring only on a narrow range of host species, mostly limited to the genus *Wisteria*, should not be taken lithely. Attempts of artificial mechanical inoculations have indeed shown that WVMV, like many other potyviruses, can infect a vast assortment of herbaceous hosts [1,2]. There was recent evidence that mixed colonies of *Aphis craccivora* and *A. fabae* colonizing a healthy locust tree (*Robinia* spp.) could vector WVMV and transmit the virus, under controlled conditions, to bean (*Phaseolus vulgaris*) seedlings [16]. We have no information on the frequency and distribution of WVMV in different crops of economic importance, such as legumes, apart from the two isolates found in soybean and jack bean, whose taxonomic attribution remains to be confirmed. However, it is not difficult to assume that a wide range of weedy species could act as reservoirs and facilitate WVMV spread and out-of-season survival [12].

5. Conclusions

In the present work, we provided unequivocal evidence for the occurrence of WVMV in Italy on *W. sinensis*. Our findings, for the first time based on molecular evidence, confirm the clues to the presence of the virus in the country in the study of Conti and Lovisolo more than 50 years ago.

Our report raises alert and calls for further efforts to monitor WVMV introduction and to prevent the spread of a serious threat to the ornamental nursery sector due to the large popularity of *W. sinensis* as an ornamental species in Europe. The circulation of WVMV-infected *W. sinensis* plants for gardening purposes, through wholesale and retail nurseries, could entail the risk of spreading the virus to different countries. It has been proved that WVMV can be disseminated by a range of vegetative propagation paths, e.g., grafting, cuttings, etc., but it is also easily transmitted by aphids in a non-persistent manner. Due to the perennial habit of wisteria, planting WVMV-infected plants in an area that is still virus-free could potentially provide a continuous inoculum source for aphid vectors, thus facilitating the extension of the WVMV host range to other susceptible species in the neighbourhood.

Our study, to the best of our knowledge, is also the first to attempt, based on the capsid protein gene, an analysis of the genetic and haplotype diversity of WVMV and estimate its population dynamics and evolutionary traits. We encourage future sequencing efforts of other isolates from different geographic areas and hosts to gain further insight into the intriguing evolutionary features of WVMV and corroborate the hypotheses we have advanced in this study.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriculture13051090/s1, Figure S1: Sequence polymorphism within WVMV haplotypes; Table S1: List of virus isolates used for sequence comparison and phylogenetic analysis in the NIb gene; Table S2: List of virus isolates used for sequence comparison, phylogenetic, genetic diversity and population dynamics analyses, in the CP gene; Table S3: Distance matrix of pairwise MAFFT-based comparison of the partial sequence (116 aa) of the NIb protein of 147 isolates representative of 143 species members of the genus *Potyvirus*; Table S4: Percent amino acid sequence identities of the NIb protein, between WVMV Bari and isolates representative of 143 potyvirus species; Table S5: Distance matrix of pairwise MAFFT-based comparison of the complete sequence of the CP of 71 isolates representative of 23 species in the genus *Potyvirus*; Table S6: Percent amino acid sequence identities of the capsid protein, between WVMV Bari and 70 isolates representative of 23 potyvirus species; Table S7: Haplotypes identified among WVMV isolates.

Author Contributions: Conceptualization, G.D., A.M., P.S. and M.M.; methodology, G.D. and M.M.; software, M.M.; validation, G.D. and M.M.; formal analysis, G.D. and M.M.; investigation, G.D. and M.M.; resources, A.M. and P.S.; data curation, M.M.; writing—original draft preparation, M.M.; writing—review and editing, G.D., A.M. and P.S.; visualization, M.M.; supervision, A.M. and P.S.; project administration, A.M. and P.S.; funding acquisition, A.M. and P.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable, as this study is not involving humans or animals.

Data Availability Statement: Nucleic acid sequence data relative to WVMV Bari isolate have been deposited in GenBank under the accession numbers OP381183.1 and OP748400.3. Any other raw dataset generated and/or analysed for the aim of the current study is available from the corresponding author upon reasonable request.

Acknowledgments: The authors are grateful to Yoonsoo Hahn, Chung-Ang University, Seoul, South Korea, for fruitful discussions on the diversity of cleavage sites in potyviruses.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Clover, G.; Tang, Z.; Smales, T.; Pearson, M. Taxonomy of Wisteria vein mosaic virus and extensions to its host range and geographical distribution. *Plant Pathol.* **2003**, *52*, 92–96. [CrossRef]
2. Al-Jaberi, M.S.; Moradi, Z.; Mehrvar, M.; Al-Inizi, H.R.; Zakiaghl, M. Whole genome characterization of Wisteria vein mosaic virus from Iran and its relationship to other members of bean common mosaic virus group. *3 Biotech* **2021**, *11*, 407. [CrossRef] [PubMed]
3. Liang, W.; Song, L.; Tian, G.; Li, H.; Fan, Z. The genomic sequence of Wisteria vein mosaic virus and its similarities with other potyviruses. *Arch. Virol.* **2006**, *151*, 2311–2319. [CrossRef]
4. Clover, G.; Denton, J.; Denton, G. First report of Wisteria vein mosaic virus on *Wisteria* spp. in the United Kingdom. *New Dis. Rep.* **2015**, *31*, 1. [CrossRef]
5. Brierley, P.; Lorentz, P. Wisteria mosaic and peony leaf curl, two diseases of ornamental plants caused by viruses transmissible by grafting but not by sap inoculation. *Plant Dis. Reptr.* **1957**, *41*, 691–693.
6. Bos, L. The identification of three new viruses isolated from *Wisteria* and *Pisum* in the Netherlands, and the problem of variation within the potato virus Y group. *Neth. J. Plant Pathol.* **1970**, *76*, 8–46. [CrossRef]
7. Al-Jaberi, M.; Mehrvar, M.; Zakiahgl, M. Molecular identification of an isolate of Wisteria vein mosaic virus in Khorasan Razavi province. *J. Plant Prot. Mashad* **2019**, *33*, 23–26.
8. Liang, W.; Song, L.; Li, Y.; Tian, G.; Li, H.; Fan, Z. First report of Wisteria vein mosaic virus in China. *Plant Pathol.* **2004**, *53*, 516. [CrossRef]
9. Zhu, F.; Zhu, P.-X.; Xu, F.; Ji, Z.-L. First report of Wisteria vein mosaic virus infecting Chinese wisteria in Jiangsu Province in China. *J. Plant Dis. Prot.* **2019**, *126*, 373–377. [CrossRef]
10. Ji, Z.-L.; Zhu, P.-X.; Ji, Y.-H.; Xu, F.; Zhu, F. First report of Wisteria vein mosaic virus in Chinese wisteria in Jiangxi Province in China. *J. Plant Pathol.* **2019**, *101*, 1259–1260. [CrossRef]
11. Kamińska, M.; Malinowski, T.; Rudzińska-Langwald, A.; Diaz, L. The occurrence of Wisteria vein mosaic virus in Wisteria floribunda DC plants in Poland. *J. Phytopathol.* **2006**, *154*, 414–417. [CrossRef]
12. Naidu, R.A.; Karthikeyan, G. First report of Wisteria vein mosaic virus in *Wisteria sinensis* in the United States of America. *Plant Health Prog.* **2008**, *9*, 42. [CrossRef]
13. Ward, L.; Tang, J.; Clover, G. First report of Wisteria vein mosaic virus on *Wisteria sinensis* in New Zealand. *Plant Dis.* **2008**, *92*, 1134. [CrossRef] [PubMed]
14. Conti, M.; Lovisolo, O. Observations on a virus isolated from *Wisteria floribunda* DC in Italy. *Riv. Patol. Veg.* **1969**, *5*, 115–132.
15. Brčák, J. A Prague isolate of Wisteria vein mosaic virus. *Biol. Plant.* **1980**, *22*, 465–469. [CrossRef]
16. Valouzi, H.; Hashemi, S.-S.; Wylie, S.J.; Ahadiyat, A.; Golnaraghi, A. Wisteria vein mosaic virus detected for the first time in Iran from an unknown host by analysis of aphid vectors. *Plant Pathol. J.* **2020**, *36*, 87–97. [CrossRef]
17. Bos, L. Wisteria vein mosaic potyvirus. In *Viruses of Plants: Descriptions and Lists from the VIDE Database*; Brunt, A.A., Ed.; CAB International: Wallingford, UK, 1996; pp. 1407–1409.
18. Yao, M.-C.; Zhu, P.-X.; Zhang, Q.-Q.; Zhang, Q.-P.; Ji, Z.-L.; Zhu, F. Development of a reverse transcription-loop mediated isothermal amplification assay for detection of Wisteria vein mosaic virus. *Eur. J. Plant Pathol.* **2022**, *163*, 113–123. [CrossRef]
19. Valder, P. *Wisterias. A Comprehensive Guide*; Florilegium: Balmain, Australia, 1995.
20. Keskin, S.; Sirin, Y.; Cakir, H.E.; Keskin, M. Phenolic composition and antioxidant properties of *Wisteria sinensis*. *Int. J. Sci. Technol. Res.* **2019**, *5*, 98–103.
21. Li, Y.; Deng, C.; Qiao, Y.; Zhao, X.; Zhou, Q. Characterization of a new badnavirus from *Wisteria sinensis*. *Arch. Virol.* **2017**, *162*, 2125–2129. [CrossRef]
22. Mehrvar, M.; Moradi, Z.; Al-Jaberi, M. First report of Wisteria badnavirus 1 infecting *Wisteria sinensis* in Iran. *New Dis. Rep.* **2022**, *46*, 2. [CrossRef]
23. Dabiri, S.; Moradi, Z.; Mehrvar, M.; Zakiaghl, M. Analysis of the complete genome sequence of cucumber mosaic virus from *Vinca minor* and *Wisteria sinensis* in Iran. *J. Plant Pathol.* **2020**, *102*, 1263–1268. [CrossRef]
24. Milojević, K.; Radović, N.; Stanković, I.; Vučurović, A.; Nikolić, D.; Bulajić, A.; Krstić, B. First report of cucumber mosaic virus infecting *Wisteria sinensis* in Serbia. *Plant Dis.* **2016**, *100*, 1799. [CrossRef]
25. Moradi, Z.; Mehrvar, M. Whole-genome characterization of alfalfa mosaic virus obtained from metagenomic analysis of *Vinca minor* and *Wisteria sinensis* in Iran with implications for the genetic structure of the virus. *Plant Pathol. J.* **2021**, *37*, 619–631. [CrossRef] [PubMed]
26. Jo, Y.; Yoon, Y.N.; Jang, Y.-W.; Choi, H.; Lee, Y.-H.; Kim, S.-M.; Choi, S.Y.; Lee, B.C.; Cho, W.K. Soybean viromes in the Republic of Korea revealed by RT-PCR and next-generation sequencing. *Microorganisms* **2020**, *8*, 1777. [CrossRef] [PubMed]
27. Jerzak, G.; Bernard, K.A.; Kramer, L.D.; Ebel, G.D. Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. *J. Gen. Virol.* **2005**, *86*, 2175. [CrossRef] [PubMed]
28. Randa-Zelyüt, F.; Fox, A.; Karanfil, A. Population genetic dynamics of southern tomato virus from Turkey. *J. Plant Pathol.* **2022**, *105*, 211–224. [CrossRef]
29. Rodríguez-Nevado, C.; Montes, N.; Pagán, I. Ecological factors affecting infection risk and population genetic diversity of a novel potyvirus in its native wild ecosystem. *Front. Plant Sci.* **2017**, *8*, 1958. [CrossRef]

30. Morelli, M.; Giampetruzzi, A.; Laghezza, L.; Catalano, L.; Savino, V.N.; Saldarelli, P. Identification and characterization of an isolate of apple green crinkle associated virus involved in a severe disease of quince (*Cydonia oblonga*, Mill.). *Arch. Virol.* **2017**, *162*, 299–306. [CrossRef]
31. Zheng, L.; Rodoni, B.; Gibbs, M.; Gibbs, A.J. A novel pair of universal primers for the detection of potyviruses. *Plant Pathol.* **2010**, *59*, 211–220. [CrossRef]
32. NCBI Virus. Available online: https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/ (accessed on 31 January 2023).
33. Marchler-Bauer, A.; Bryant, S.H. CD-Search: Protein domain annotations on the fly. *Nucleic Acids Res.* **2004**, *32*, W327–W331. [CrossRef]
34. Hall, T.A. *BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT*; Nucleic Acids Symposium Series; Information Retrieval Ltd.: London, UK, 1999; pp. 95–98.
35. Wernersson, R.; Pedersen, A.G. RevTrans: Multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic acids res.* **2003**, *31*, 3537–3539. [CrossRef] [PubMed]
36. Foster, P.G.; Hickey, D.A. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *J. Mol. Evol.* **1999**, *48*, 284–290. [CrossRef] [PubMed]
37. Katoh, K.; Misawa, K.; Kuma, K.i.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [CrossRef] [PubMed]
38. Muhire, B.M.; Varsani, A.; Martin, D.P. SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE* **2014**, *9*, e108277. [CrossRef] [PubMed]
39. Lemoine, F.; Correia, D.; Lefort, V.; Doppelt-Azeroual, O.; Mareuil, F.; Cohen-Boulakia, S.; Gascuel, O. NGPhylogeny. fr: New generation phylogenetic services for non-specialists. *Nucleic Acids Res.* **2019**, *47*, W260–W265. [CrossRef] [PubMed]
40. Desper, R.; Gascuel, O. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. In Proceedings of the Algorithms in Bioinformatics: Second International Workshop, WABI 2002, Rome, Italy, 17–21 September 2002; pp. 357–374.
41. Lefort, V.; Desper, R.; Gascuel, O. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* **2015**, *32*, 2798–2800. [CrossRef]
42. Criscuolo, A.; Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **2010**, *10*, 210. [CrossRef]
43. Junier, T.; Zdobnov, E.M. The Newick utilities: High-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* **2010**, *26*, 1669–1670. [CrossRef]
44. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]
45. Goh, C.J.; Hahn, Y. Analysis of proteolytic processing sites in potyvirus polyproteins revealed differential amino acid preferences of NIa-Pro protease in each of seven cleavage sites. *PLoS ONE* **2021**, *16*, e0245853. [CrossRef]
46. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.E.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Springer Humana Press: Totowa, NJ, USA, 2005.
47. Martin, D.P.; Murrell, B.; Golden, M.; Khoosal, A.; Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **2015**, *1*, vev003. [CrossRef] [PubMed]
48. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [CrossRef] [PubMed]
49. Jukes, T.H.; Cantor, C.R. Evolution of protein molecules. In *Mammalian Protein Metabolism*; Munro, H.N., Ed.; Academic Press: New York, USA, 1969; Volume 3, pp. 21–132.
50. Nei, M.; Miller, J.C. A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics* **1990**, *125*, 873–879. [CrossRef] [PubMed]
51. Okonechnikov, K.; Golosova, O.; Fursov, M.; Team, U. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [CrossRef]
52. Wright, S. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **1965**, *19*, 395–420. [CrossRef]
53. Garg, S.; Alam, M.T.; Das, M.K.; Dev, V.; Kumar, A.; Dash, A.P.; Sharma, Y.D. Sequence diversity and natural selection at domain I of the apical membrane antigen 1 among Indian *Plasmodium falciparum* populations. *Malar. J.* **2007**, *6*, 154. [CrossRef]
54. Slatkin, M. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **1993**, *47*, 264–279. [CrossRef]
55. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595. [CrossRef]
56. Fu, Y.-X.; Li, W.-H. Statistical tests of neutrality of mutations. *Genetics* **1993**, *133*, 693–709. [CrossRef]
57. Shen, W.; Shi, Y.; Dai, Z.; Wang, A. The RNA-dependent RNA polymerase NIb of potyviruses plays multifunctional, contrasting roles during viral infection. *Viruses* **2020**, *12*, 77. [CrossRef]
58. Ferrer-Orta, C.; Ferrero, D.; Verdaguer, N. RNA-dependent RNA polymerases of picornaviruses: From the structure to regulatory mechanisms. *Viruses* **2015**, *7*, 4438–4460. [CrossRef] [PubMed]

59. Revers, F.; Le Gall, O.; Candresse, T.; Maule, A.J. New advances in understanding the molecular biology of plant/potyvirus interactions. *Mol. Plant-Microbe Interact.* **1999**, *12*, 367–376. [CrossRef]
60. Lopez-Moya, J.; Wang, R.; Pirone, T. Context of the coat protein DAG motif affects potyvirus transmissibility by aphids. *J. Gen. Virol.* **1999**, *80*, 3281–3288. [CrossRef]
61. Zhang, B.; Li, Q.; Hu, J.; Zhang, L.; Dong, X.; Ji, P.; Dong, J. Complete genome sequence analysis of a new potyvirus isolated from *Paris polyphylla* var. yunnanensis. *Arch. Virol.* **2023**, *168*, 43. [CrossRef] [PubMed]
62. Atreya, P.L.; Atreya, C.D.; Pirone, T.P. Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7887–7891. [CrossRef]
63. Adams, M.; Antoniw, J.; Fauquet, C. Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Arch. Virol.* **2005**, *150*, 459–479. [CrossRef]
64. ISO. *ISO 3166-1: Codes for the Representation of Names of Countries and Their Subdivisions–Part 1: Country Codes*; ISO: Geneva, Switzerland, 2006.
65. De Jong, M.A.; Wahlberg, N.; Van Eijk, M.; Brakefield, P.M.; Zwaan, B.J. Mitochondrial DNA signature for range-wide populations of Bicyclus anynana suggests a rapid expansion from recent refugia. *PLoS ONE* **2011**, *6*, e21385. [CrossRef]
66. Zheng, L.; Wayper, P.J.; Gibbs, A.J.; Fourment, M.; Rodoni, B.C.; Gibbs, M.J. Accumulating variation at conserved sites in potyvirus genomes is driven by species discovery and affects degenerate primer design. *PLoS ONE* **2008**, *3*, e1586. [CrossRef]
67. Joshi, V. New/unusual or noteworthy plant disease diagnoses—1997. In *Crop Protection Newsletter 20*; Canadian Ministry of Agriculture, Food and Fisheries: Kelowna, British Columbia, Canada, 1998.
68. Nakasato, K.; Fujioka, S.; Sugawara, Y.; Ono, T.; Nishio, T.; Tsuda, S. First detection of two potyviruses, Uraria mosaic virus and Passiflora mosaic virus Y, from passionfruit in Japan. *J. Gen. Plant Pathol.* **2020**, *86*, 401–404. [CrossRef]
69. Chen, C.; Chang, C.; Tsai, H.; Hsu, H. Identification of a potyvirus causing latent infection in calla lilies. *Plant Dis.* **2004**, *88*, 1046. [CrossRef]
70. Criscuolo, A.; Gascuel, O. Fast NJ-like algorithms to deal with incomplete distance matrices. *BMC Bioinform.* **2008**, *9*, 166. [CrossRef]
71. Gibbs, A.; Ohshima, K. Potyviruses and the digital revolution. *Annu. Rev. Phytopathol.* **2010**, *48*, 205–223. [CrossRef] [PubMed]
72. Gibbs, A.J.; Hajizadeh, M.; Ohshima, K.; Jones, R.A. The potyviruses: An evolutionary synthesis is emerging. *Viruses* **2020**, *12*, 132. [CrossRef] [PubMed]
73. Herath, V.; Romay, G.; Urrutia, C.D.; Verchot, J. Family level phylogenies reveal relationships of plant viruses within the Order Bunyavirales. *Viruses* **2020**, *12*, 1010. [CrossRef] [PubMed]
74. Gibbs, A.J.; Ohshima, K.; Phillips, M.J.; Gibbs, M.J. The prehistory of potyviruses: Their initial radiation was during the dawn of agriculture. *PLoS ONE* **2008**, *3*, e2523. [CrossRef]
75. Gibbs, A.J.; Trueman, J.; Gibbs, M.J. The bean common mosaic virus lineage of potyviruses: Where did it arise and when? *Arch. Virol.* **2008**, *153*, 2177–2187. [CrossRef]
76. Worrall, E.A.; Wamonje, F.O.; Mukeshimana, G.; Harvey, J.J.; Carr, J.P.; Mitter, N. Bean common mosaic virus and bean common mosaic necrosis virus: Relationships, biology, and prospects for control. *Adv. Virus Res.* **2015**, *93*, 1–46.
77. Yasaka, R.; Fukagawa, H.; Ikematsu, M.; Soda, H.; Korkmaz, S.; Golnaraghi, A.; Katis, N.; Ho, S.Y.; Gibbs, A.J.; Ohshima, K. The timescale of emergence and spread of turnip mosaic potyvirus. *Sci. Rep.* **2017**, *7*, 4240. [CrossRef]
78. Gao, F.; Zou, W.; Xie, L.; Zhan, J. Adaptive evolution and demographic history contribute to the divergent population genetic structure of potato virus Y between China and Japan. *Evol. Appl.* **2017**, *10*, 379–390. [CrossRef]
79. Nigam, D.; LaTourrette, K.; Souza, P.F.; Garcia-Ruiz, H. Genome-wide variation in potyviruses. *Front. Plant Sci.* **2019**, *10*, 1439. [CrossRef]
80. Hughes, A.L. Small effective population sizes and rare nonsynonymous variants in potyviruses. *Virology* **2009**, *393*, 127–134. [CrossRef] [PubMed]
81. Nei, M. *Molecular Evolutionary Genetics*; Columbia University Press: New York, NY, USA, 1987.
82. Hughes, A.L. *Adaptive Evolution of Genes and Genomes*; Oxford University Press: New York, NY, USA, 1999.
83. Moradi, Z.; Mehrvar, M. Metagenomic Analysis of *Malva sylvestris* from Iran displays a Malva vein clearing virus genome. *J. Agric. Sci. Technol.* **2023**, *25*, 213–223. [CrossRef]
84. García-Arenal, F.; Fraile, A.; Malpica, J.M. Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* **2001**, *39*, 157–186. [CrossRef] [PubMed]
85. Valouzi, H.; Shahmohammadi, N.; Golnaraghi, A.; Moosavi, M.R.; Ohshima, K. Genetic diversity and evolutionary analyses of potyviruses infecting narcissus in Iran. *J. Plant Pathol.* **2022**, *104*, 237–250. [CrossRef] [PubMed]
86. Sokhandan-Bashir, N.; Melcher, U. Population genetic analysis of grapevine fanleaf virus. *Arch. Virol.* **2012**, *157*, 1919–1929. [CrossRef]
87. Jones, D.R. Plant viruses transmitted by thrips. *Eur. J. Plant Pathol.* **2005**, *113*, 119–157. [CrossRef]
88. Bashalkhanov, S.; Pandey, M.; Rajora, O.P. A simple method for estimating genetic diversity in large populations from finite sample sizes. *BMC Genet.* **2009**, *10*, 84. [CrossRef]
89. Carbonell, A.; Maliogka, V.I.; Pérez, J.d.J.; Salvador, B.; León, D.S.; García, J.A.; Simón-Mateo, C. Diverse amino acid changes at specific positions in the N-terminal region of the coat protein allow Plum pox virus to adapt to new hosts. *Mol. Plant-Microbe Interact.* **2013**, *26*, 1211–1224. [CrossRef]

90. Decroocq, V.; Salvador, B.; Sicard, O.; Glasa, M.; Cosson, P.; Svanella-Dumas, L.; Revers, F.; García, J.; Candresse, T. The determinant of potyvirus ability to overcome the RTM resistance of *Arabidopsis thaliana* maps to the N-terminal region of the coat protein. *Mol. Plant-Microbe Interact.* **2009**, *22*, 1302–1311. [CrossRef]
91. Moury, B.; Simon, V. dN/dS-based methods detect positive selection linked to trade-offs between different fitness traits in the coat protein of potato virus Y. *Mol. Biol. Evol.* **2011**, *28*, 2707–2717. [CrossRef]
92. Ivanov, K.; Eskelin, K.; Lohmus, A.; Mäkinen, K. Molecular and cellular mechanisms underlying potyvirus infection. *J. Gen. Virol.* **2014**, *95*, 1415–1429. [CrossRef]
93. Nyakaana, S.; Tumusiime, C.; Oguge, N.; Siegismund, H.R.; Arctander, P.; Muwanika, V. Mitochondrial DNA diversity and population structure of a forest-dependent rodent, *Praomys taitae* (Rodentia: Muridae) Heller 1911, in the fragmented forest patches of Taita Hills, Kenya. *S. Afr. J. Sci.* **2008**, *104*, 499–504. [CrossRef]
94. Li, Y.; Haseneyer, G.; Schön, C.-C.; Ankerst, D.; Korzun, V.; Wilde, P.; Bauer, E. High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (*Secale cereale* L.) genes involved in frost response. *BMC Plant Biol.* **2011**, *11*, 6. [CrossRef] [PubMed]
95. Slatkin, M.; Barton, N.H. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **1989**, *43*, 1349–1368. [CrossRef] [PubMed]
96. Nagy, P.D. Recombination in plant RNA viruses. In *Plant Virus Evolution*; Roossinck, M.J., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 133–156.
97. van der Walt, E.; Rybicki, E.P.; Varsani, A.; Polston, J.; Billharz, R.; Donaldson, L.; Monjane, A.L.; Martin, D.P. Rapid host adaptation by extensive recombination. *J. Gen. Virol.* **2009**, *90*, 734. [CrossRef] [PubMed]
98. Muhammad, K.; Herath, V.; Ahmed, K.; Tahir, M.; Verchot, J. Genetic diversity and molecular evolution of sugarcane mosaic virus, comparing whole genome and coat protein sequence phylogenies. *Arch. Virol.* **2022**, *167*, 2239–2247. [CrossRef]