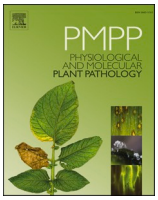
Physiological and Molecular Plant Pathology 125 (2023) 102008

Contents lists available at ScienceDirect

Physiological and Molecular Plant Pathology

journal homepage: www.elsevier.com/locate/pmpp



# Population structure and genetic diversity of tobacco mild green mosaic virus variants in Western Anatolia of Turkey

Ali Karanfil [a], Filiz Randa-Zelyüt [b,c], Savaş Korkmaz [a,*]

[a] Department of Plant Protection, Faculty of Agriculture, Canakkale Onsekiz Mart University, Canakkale, Turkey
[b] Department of Plant Protection, Faculty of Agriculture and Natural Sciences, Bilecik Seyh Edebali University, Bilecik, Turkey
[c] Biotechnology Application and Research Centre, Bilecik Şeyh Edebali University, 11230 Gulumbe Campus, Bilecik, Turkey

ARTICLE INFO

*Keywords:*
Tobacco
Tobamovirus
Genetic diversity
Selection pressure

ABSTRACT

Tobamoviruses have been pioneers in understanding the population genetic structure, host-vector interactions, and evolutionary processes of viruses. To better understand the population genetic structure and molecular evolutionary relationships of tobacco mild green mosaic virus (TMGMV), a significant species of tobamoviruses, extensive analyses were performed using bioinformatics tools in this study. 300 samples were collected from plants exhibiting viruses and virus-like symptoms from Turkey's largest tobacco cultivation areas during the plant vegetation period between 2019 and 2020. Samples were tested using conventional molecular techniques for tobamoviruses, including ToBRFV, TMV, TMGMV, ToMV, and ToMMV species. Single and double tobamovirus infections were determined in 258 of 300 samples. Single infections were 225 (75%), 10 (3.33%), and 1 (0.33%) for TMGMV, TMV, and ToMV, respectively. Double infections were 21 (7.0%) and 1 (0.33%) for TMGMV + TMV and TMV + ToMV, respectively. TMGMV-infected samples from each region were chosen, and their complete genomes were revealed. Detailed analyses were carried out from Turkey, with global variants available in GenBank. Molecular evolutionary analyses revealed three main lineages (Clades I, II, and III) at the p126, p183, MP, and complete genome levels and two main lineages at the CP gene (Clades I and II). Variants from Turkey were distributed in two different branches in other phylogenetic trees, except for the CP gene. High haplotype diversity and low nucleotide diversity were determined in each gene region, indicating consistent genetic stability. In addition, purifying selection pressures were determined in specific gene regions. With AMOVA (analysis of molecular variance), it was determined that the source of genetic variation came from within the main lineages, including various hosts and geographies. The differences in lineages were confirmed by independent test statistics. While neutrality tests revealed population expansions in the CP and MP genes, other p183 and p126 genes revealed bottlenecks or balancing selection. The fact that TMGMV was more common than TMV in Turkey strongly supported the phenomenon called "mutational melting" or "Müller latch," which presumably causes TMV to disappear from its niche.

## 1. Introduction

Tobacco (*Nicotiana tabacum*), a member of the Solanaceae family, is cultivated throughout many temperate regions of the world. In Turkey, tobacco farming is more common in Western Anatolia, which consists of the Aegean and Marmara regions. Therefore, 90% of the total tobacco production in the country is obtained from these regions [1], providing a significant level of agricultural employment in rural areas [2]. Turkish or Oriental tobacco maintains its strategic importance in terms of the economy as it is frequently preferred in the global market for its unique aroma and quality [3].

Several diseases caused by phytopathogens, which adversely affect the quantity and quality of tobacco plants, have been reported. However, viral diseases, primarily caused by different species of the genus Tobamovirus, have been identified as one of the most important factors limiting tobacco cultivation worldwide. Furthermore, tobamoviruses, one of the genera of the *Virgaviridae* family, have a distinctive feature from the other genera of the family because they have an undivided genome [4]. The virions of the members of this genus have a length of 300–310 nm and a diameter of 18 nm, and their genomes feature positive polarity (+) ssRNA and are 6.3–6.6 kb in size [4,5]. The current 37 species of tobamovirus can be further divided into eight groups based on

* Corresponding author.
*E-mail addresses:* ali.karanfil@hotmail.com (A. Karanfil), filizrandazelyut@gmail.com (F. Randa-Zelyüt), skorkmaz@comu.edu.tr (S. Korkmaz).

https://doi.org/10.1016/j.pmpp.2023.102008
Received 24 February 2023; Received in revised form 15 March 2023; Accepted 21 March 2023
Available online 22 March 2023

their sequence, genome organization, antigenic relationships, and host range. Moreover, these groups, which are specific to the host, are classified according to the plant families: Solanaceae, Brassicaceae, Cucurbitaceae, Malvaceae, Cactaceae, Passifloraceae, Apocynaceae, and Leguminosae [6–11].

Tobacco mosaic virus (TMV), tomato brown rugose fruit virus (ToBRFV), tomato mottle mosaic virus (ToMMV), and tomato mosaic virus (ToMV) are two of the most well-known species of a group particularly specialized in infecting members of Solanaceae. In addition to this group, another member is the tobacco mild green mosaic virus (TMGMV) [12]. However, it has been reported from different countries that TMGMV infects various hosts, including vegetables, industrial, and ornamental plants [13–16]. The presence of the virus in Turkey was reported for the first time in the Manisa province which is located in the Aegean region, which has a prevalent tobacco plantation [17].

Four proteins are known to be encoded by the four open reading frames (ORFs) of the tobamovirus genome, and detailed studies on the roles of these proteins have primarily been conducted on the type-member TMV. The two proteins encoded by ORF1 and ORF2, with molecular masses of 124–132 kDa (p126 gene) and 181–189 kDa (p183 gene), respectively, which are responsible for the replication of the viral genome, are essential common genomic features of the genus [18]. However, the methyltransferase, helicase, and unconserved region II (NON-II) multiple domains of the p126 protein of TMV have host RNA silencing repressor-specific mechanisms [19]. Furthermore, p126/183 silencing repressor proteins have an essential role in the transmission of viral particles from cell to cell or over long distances, and they are also effective in host metabolism change to facilitate replication of the viral genome [20,21]. The remaining two ORFs, ORF3 and ORF4, encode movement protein (MP) and coat protein (CP), which have molecular weights of 28–31 kDa and 17–18 kDa, respectively [9].

Analytical methods have been developed based on genome analyses to statistically predict the impact of different evolutionary forces and mechanisms shaping plant virus populations [22]. These methods are used more frequently today, and the findings provide insight into important issues such as the epidemiology of viruses, the effects of virus-vector-host interactions on the viral genome, and how these are reflected in genetic diversity [22,23]. More specifically, tobamoviruses have been pioneers in genetic diversity research, and the investigation of TMV and tobamovirus species at the genetic level has greatly contributed to the development of all areas of virology, including the evolution of viruses [24]. However, studies on the genetic structure and diversity of TMGMV are very limited. High genetic stability was found in the variants derived from the wild plant *Nicotiana glauca* Grah. which was naturally infected with the virus during the studies conducted to determine the genetic structure of the TMGMV agent [25,26]. Another study reported that low genetic diversity may be due to different factors including positive or negative selection or recent adaptation of new environments or host plants [27].

Tobamoviruses are an ancient genus, as indicated by the fact that they share the same age as and have undergone co-evolution with their angiosperm hosts, which are thought to be 120–140 million years old [12]. Within the scope of this study, the population dynamics of TMGMV, one of the tobamoviruses that have made a great contribution to virus evolution studies, were researched by using genetic approaches. Evolutionary aspects of the virus, evolutionary forces shaping its population structure, and genetic diversity were comprehensively investigated by using gene sequences of the TMGMV reported from different geographies in GenBank, especially virus variants obtained from tobacco-growing areas of Turkey. Thus, population genetic analyses were performed at levels of the complete genome and four genes, including CP, MP, p126, and p183 of the virus.

## 2. Materials and methods

### 2.1. Field surveys

Between 2019 and 2020, field surveys were conducted in two provinces of the Marmara region (Çanakkale and Balıkesir) and five provinces of the Aegean region (Manisa, İzmir, Uşak, Aydın, and Denizli), which are the main tobacco-growing areas of Western Anatolia. Tobacco fields were randomly selected, and samples were taken from plants exhibiting virus and/or virus-like symptoms. In addition, the number of tobacco samples was determined in proportion to the production amount of each province (Fig. 1 and Supp. Table 1).

### 2.2. Tobamoviruses detection

Tobamoviruses including the ToBRFV, ToMMV, TMGMV, TMV, and ToMV species were molecularly identified at the species level using reverse transcription polymerase chain reaction (RT-PCR) studies with gene-specific primer pairs. For this purpose, total nucleic acid (TNA) was isolated from the leaves using the CTAB method with minor modifications described by Li et al. [28]. To produce cDNA libraries from TNAs, syntheses were first performed using the cDNA Synthesis Kit (Takara, Japan) and a random hexamer primer (5′–NNNNNN–3′). These generated cDNA libraries were utilized in amplifications with the 2X Emerald PCR Master Mix (Takara, Japan) and the gene-specific primer pairs listed in Supp. Table 2. The reaction results were checked under a UV imager on agarose gel electrophoresis stained with 1.5% EtBr.

### 2.3. Complete genome analyses of tobacco mild green mosaic virus

At least one TMGMV isolate from each province was chosen for complete genome analysis. Using the TNAs obtained in the virus detection studies, the six fragments of the TMGMV full genome were amplified by RT -PCR (Supp. Fig. 1) using primer pairs in Supp. Table 3. The Omega-Bio PCR Purification Kit (USA) was used to purify each obtained fragment. The purified PCR products were cloned using the T-A cloning kit from Promega (USA), and one plasmid was sequenced bidirectionally by sanger sequencing method by using M13F-R universal primer pair for each gene region. The obtained raw sequence data were assembled with the program CLC Main Workbench V.20.3 and then deposited in GenBank (OK149275-84) (Supp. Table 4).

The sequences of other isolates already in the GenBank were aligned with the obtained isolates to determine the nucleotide (nt) and amino acids (aa) similarities among them, using Sequence Demarcation Tool V.1.2 [29] for respective analysis (Supp. Table 5).

### 2.4. Phylogenetic analyses

All TMGMV isolates from Turkey and other geographies were used in phylogenetic analyses to reveal the molecular evolutionary relationships of the complete genome of the virus and each gene region. Analyses were performed in the MEGA 11 software [30]. The nucleotide sequences of variants obtained from Turkey and other global isolates in GenBank were aligned in MEGA 11 using the ClustalW parameter [31].

The nt substitution models were selected according to the lowest scores on the Bayesian Information Criterion (BIC) in MEGA 11 [30,32]. The phylogenetic tree of 98CP gene sequences was generated by the Minimum Evolution (ME) method using the Tamura-3 parameter with gamma distribution (T92 + G) and the Neighbor-Joining (NJ) algorithm; an outgroup was not used [33–35]. Moreover, evolutionary trees for the p126, p183, and MP gene regions and the whole genome of TMGMV were constructed using the Tamura-Nei (TN93) parameter and the NJ method. Four tobamovirus species; ToMV, ToMMV, TMV, and ToBRFV were used as outgroups in constructing these trees. Confidence values of isolate clusters were tested with 1000 bootstrap replicates [36].

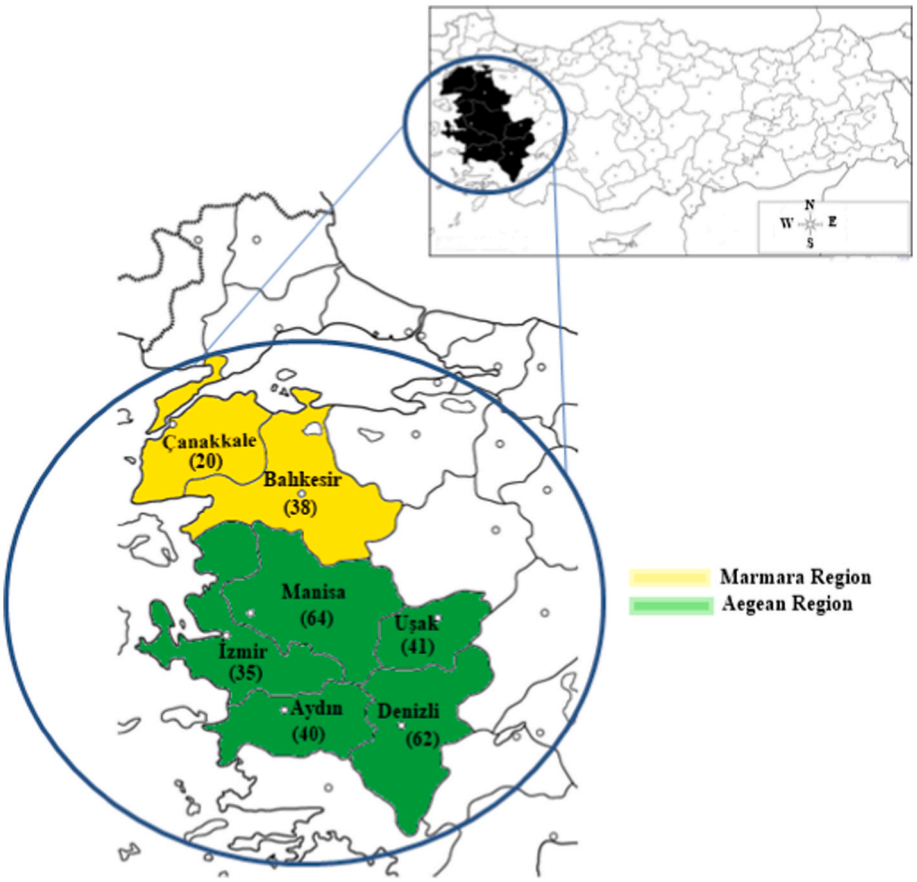



**Fig. 1.** Areas of virus surveys for sampling in Turkey (Each color symbolizes a geographic area in Turkey, and the numbers in parentheses correspond to the sampling numbers in each province).

### 2.5. Recombination and population structure analyses

To evaluate RNA recombination, which is one of the significant forces driving genetic diversity, analysis of the complete genome of TMGMV was implemented in the RDP5 package using RDP [37], BootScan [38], GENECONV [39], MaxChi [40], Chimaera [41], and SiScan [42], recombination detection programs [43]. For the detection of possible recombinant isolates, multiple alignments of the nucleotide sequence in the complete protein-coding genome region of the TMGMV isolates aligned by the program ClustalW in the CLC Main Workbench version 20.3 were transferred to RDP5 [43] analyzed for recombination.

The effects of haplotype distributions of variants used in molecular evolutionary studies on total variation were determined by molecular analysis of variance (AMOVA) [44]. Analyses were performed in Arlequin version 3.5.2.2 [45] to assess the distribution of genetic diversity. The total variance was divided into its components due to within- and between-phylogroups variance. The *F*-statistic was used to estimate the genetic differentiation between TMGMV phylogroups, and a nonparametric permutation analysis based on *10,000* repetitions was utilized to determine its significance [45,46]. Thus, data sets containing each gene region (CP, MP, p126, and p183) and a complete genome of the TMGMV were prepared based on the haplotypes.

### 2.6. Genetic diversity, neutrality, and genetic differentiation/gene flow analyses

Analyses of genetic diversity and differentiation, gene flow, and neutrality tests were performed using DnaSP v.6.12.03 software based on genes of p183, p126, MP, CP, and the complete genome of TMGMV [47]. The average pairwise nucleotide diversity ($\pi$) [48], haplotype diversity ($H_d$), haplotypes ($H$), and the ratio of non-synonymous and synonymous nt diversity $\omega$ ($d_N/d_S$) ratio were calculated according to phylogenetic groups. The gene is under negative (purifying), neutral, or positive (diversifying) selection when $\omega$ ratio is $< 1$, $= 1$ and, $>1$, respectively [47]. Furthermore, to assess natural selection among the variations of phylogenetic groups, Tajima's *D* and Fu and Li's *D*& F** (default window length of 100 sites and step size of 25 sites; without out-group) statistical tests were performed [49,50].

Genetic differentiation between TMGMV phylogroups was evaluated using independent permutation statistical tests, including Ks*, Z*, and $S_{nn}$ with *1000* replicates [51,52]. In addition, if the test statistics had a strong level of support (*P* values $< 0.05$), the null hypothesis that there was no genetic differentiation was rejected. Furthermore, the degree of gene flow among populations composed of phylogroups of TMGMV was estimated from values derived from the standardized variance of allele frequencies (Fixationindex-$F_{ST}$) and migration rate ($N_m$). If $F_{ST} > 0.33$ or $Nm < 1$, it is likely that infrequent gene flow has occurred, whereas if $F_{ST} < 0.33$ or $Nm > 1$, frequent gene flow is likely to have occurred.

## 3. Results

### 3.1. Field observations and tobamovirus distributions

During the field studies, a total of 300 plant samples exhibiting typical virus and virus-like symptoms were collected from the Marmara and Aegean regions, which are the main tobacco-producing areas of Western Anatolia. Mild and/or severe mosaic symptoms were observed in almost all samples. In addition to these symptoms, stunting was also found in very few plants.

Single and mixed virus infections were found in 258 out of 300

samples, according to molecular identification studies. The most detected virus was TMGMV, while the least detected was ToMV. In addition, while single infections were determined in 236 tobacco plants, tobamovirus co-infections were detected in the other 22 plants in the form of binary combinations (Table 1). A mild mosaic between leaf veins was very common on samples infected by TMGMV (Fig. 2). This could be a significant symptom of the TMGMV infection in tobacco. On the other hand, ToBRFV and ToMMV agents were not detected in PCR assays.

More specifically, 86% of the samples were found to be infected by at least one tobamovirus. Single infection rates were 75%, 3.33%, and 0.33% for TMGMV, TMV, and ToMV agents, respectively. Furthermore, co-infection rates were 7.0% and 0.33% for TMGMV + TMV and TMV + ToMV, respectively. All these results revealed that TMGMV is the major virus, while TMV and ToMV are minor viruses in the distribution of tobamovirus infection in the Marmara and Aegean regions, which are the largest tobacco cultivation areas of Turkey.

### 3.2. Complete protein-coding genome characterization of tobacco mild green mosaic virus

Ten new variants in total, from each of the seven provinces surveyed, were randomly chosen for complete genome sequencing. The results revealed that the cumulative length of all protein-coding gene regions was 6.309 nucleotides for all ten isolates. In detail, the length of the p183 gene region (RNA-dependent RNA polymerase; RdRp) was 4.030 nt, the p126 gene region (helicase) was 3.336 nt, the movement protein gene region (MP) was 771 nt, and the CP gene region was 480 nt for all ten Turkish variants. Complete protein-coding genomes of the ten Turkish variants had 97–99% similarity rates among themselves and 96–98% with other global variants.

The nt similarity rates among ten Turkish variants obtained in this study were determined to be 98–100% for the CP gene, 97–99% for the MP gene, 98–99% for the p126 gene, and 96–99% for the p183 gene. Moreover, the results of sequence similarity analyses also indicated that the Turkish variants shared 97–100% nt identity in the CP gene, 97–99% in the MP gene, 96–99% in the p126 gene, and 96–98% in the p183 gene with the global variant.

Amino acid sequence similarity rates of 99–100% for the CP gene, 98–100% for the MP gene, and 97–99% for both the p126 and p183 genes were estimated between ten Turkish variants. Furthermore, the aa sequence similarity analyses also indicated that the ten Turkish variants shared 98–100% homologies in the CP gene, 92–100% in the MP gene, 95–99% in the p126 gene, and 96–99% in the p183 gene with global isolates.

### 3.3. Phylogenetic inferences

To infer molecular evolutionary relationships among TMGMV variants, aligned nt datasets generated from Turkey and other global variants were examined separately at the gene and complete genome levels. The phylogenetic tree for the CP gene was created unrooted, without using an outgroup. For protein-coding gene regions (MP, p126, and p183 genes) and the whole genome, ToMV (access no MH393622), ToBRFV (access no MK648157), ToMMV (access no KU594507), and TMV (access no AF273221) tobamoviruses were used as outgroups.

Molecular evolutionary analyses revealed that TMGMV variants formed three main clades (I, II, and III) supported by high bootstrap values (≥90) at the complete protein-coding or whole genome level (Fig. 3). Moreover, the main clades I and III were divided into two different subgroups, a and b. The main clade I contained 28 variants from China, the USA, Vietnam, Jordan, Italy, Germany, Slovenia, and Turkey (AYD-13, AYD-19, DEN-23, DEN-24, USA-46, and CAN-77). Clade II included the four variants from Kenya, Spain, Brazil, and Slovenia. Clade III contained Taiwan, Japan, and Turkey variants (MAN-1_2, MAN -5_7, BAL-6, and IZM-37). Furthermore, of the 10 variants from Turkey in the two main clades (I and III), six were in subgroup I-a and four in subgroup III-b (Fig. 3).

In the phylogenetic analyses performed according to different gene regions of TMGMV, all variants were divided into three main clades in each comparison based on the p126, p183, and MP genes. Furthermore, evolutionary tree models of these three genes closely resembled the topology of the resulting tree for the whole genome. Thus, three main clades (I, II, and III) emerged for each gene region, and two separate subgroups for clades I and II. More specifically, the distributions of Turkish and global variants in main and sub-branches were quite similar to each other, and their clustering was supported by high confidence values (≥90) (Fig. 3). Thus, for all three gene regions, six of the Turkish TMGMV variants were grouped with the variants from China, the USA, Vietnam, Jordan, Italy, Germany, and Slovenia in the main clade I, and four with the variants from Taiwan and Japan in the main clade III.

Two major clades appeared based on the complete CP gene of TMGMV in the evolutionary tree. Sixty-five of the ninety-eight variants from Turkey and other countries (France, USA, Germany, Spain, Italy, Slovenia, China, Vietnam, Taiwan, Jordan Hungary, Brazil, and South Korea) were located in clade I, and the remaining thirty-three in clade II. Thus, major clade I consisted of variants from Turkey and other geographies. With the exception of one Chinese variant, clade II was entirely composed of Spanish variants. Clade II of the CP gene has been referred to in this study as the 'Spanish Clade' as a consequence (Fig. 3).

### 3.4. Recombination and AMOVA results

The main criteria in order to consider true recombination is accepted to obtain a P-value less than 1.0 × 10–6 with at least three different methods [53]. Otherwise, no recombination was considered. And, the existing recombination events among TMGMV isolates were searched with program RDP5 with 7 recombination detection models, there was no obtain the statistical significance of the recombination signal. For this reason, it was not determined any recombinant TMGMV isolate.

Molecular analysis of variance (AMOVA) was used to assess the genetic variation distribution among and within the main clades of populations for the complete genome and specific gene regions of TMGMV based on haplotypes of phylogenetic groups. Based on the complete genome, the AMOVA values demonstrated significant differentiation

**Table 1**
The number of single and double tobamovirus infections in the collected samples.

| Region | Province | TMGMV | TMV | ToMV | TMGMV + TMV | TMV + ToMV | No of infected and collected samples |
|---|---|---|---|---|---|---|---|
| **Marmara** | Çanakkale | 4 | 5 | 1 | 2 | 1 | 13/20 |
| | Balıkesir | 12 | 2 | 0 | 11 | 0 | 25/38 |
| **Aegean** | İzmir | 32 | 0 | 0 | 0 | 0 | 32/35 |
| | Manisa | 60 | 1 | 0 | 0 | 0 | 61/64 |
| | Uşak | 21 | 2 | 0 | 7 | 0 | 30/41 |
| | Aydın | 37 | 0 | 0 | 0 | 0 | 37/40 |
| | Denizli | 59 | 0 | 0 | 1 | 0 | 60/62 |
| **Total** | | 225 (75%) | 10 (3.33%) | 1 (0.33%) | 21 (7.0%) | 1 (0.33%) | 258/300 (86%) |

**Fig. 2.** Typical mosaic symptoms on tobacco foliar areas induced by single infection of tobacco mild green mosaic virus under field conditions.

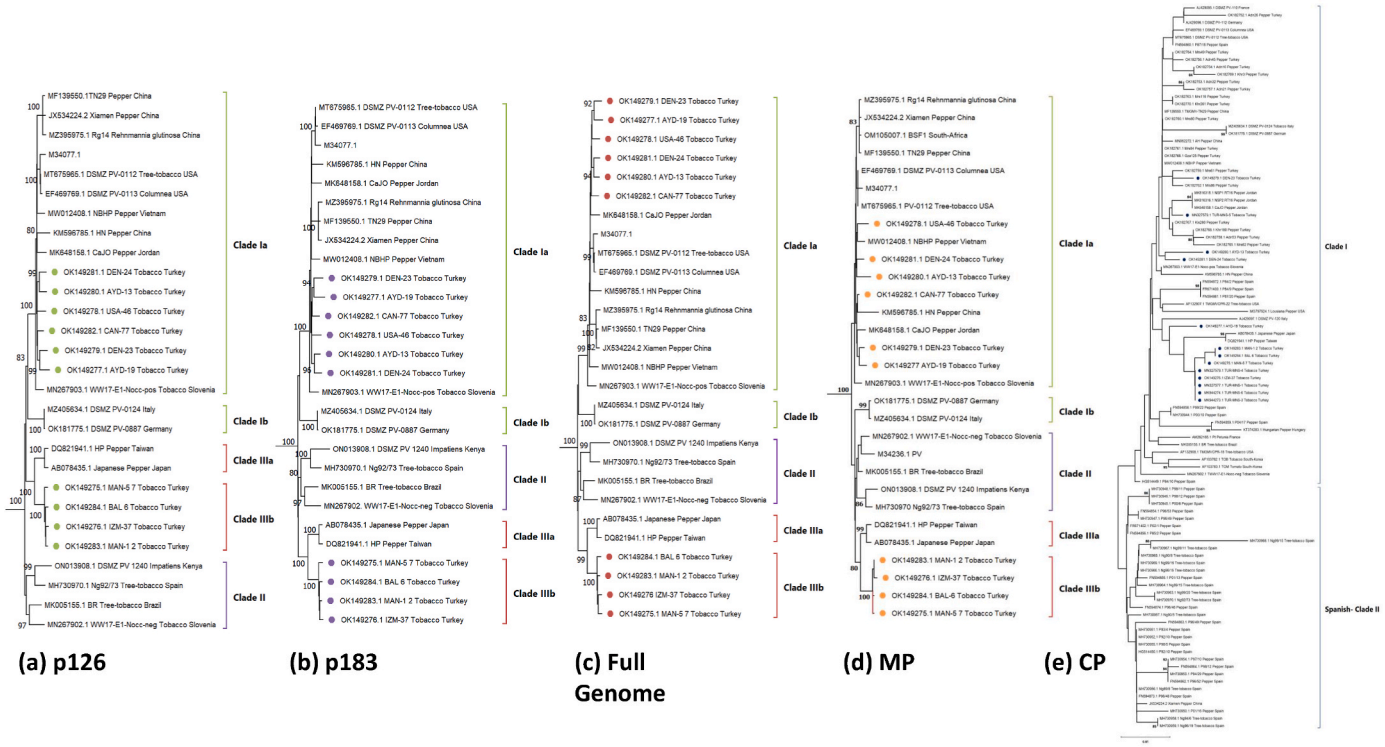



**Fig. 3.** Phylogenetic relationships of tobacco mild green mosaic virus isolates. (a, b, d) For protein-coding gene regions (p126, p183, and MP) and (c) the complete genome, ToMV (access no MH393622), ToBRFV (access no MK648157), ToMMV (access no KU594507), and TMV (access no AF273221) tobamoviruses were used as outgroups. (e) The phylogenetic tree for the CP gene was created unrooted, without using an outgroup.

between the three main clades (I, II, and III), with 46.85% of the total variation was due to differentiation among clades and 53.15% of the genetic variation was due to differentiation within clades (Table 2). In terms of genetic differentiation values between and within clade populations, the molecular variance values found for the phylogroups of the MP, p126, and p183 gene regions (I, II, and III) were very similar to those obtained from the whole genome. Additionally, genetic variation within the main clades contributed more than genetic variation among clades (42.22%), with a value of 57.78% for phylogroups of the CP gene (Table 2).

### 3.5. *Structure of the TMGMV population and the direction of selective forces*

Genetic variations were estimated using nucleotide diversity ($\pi$), haplotype diversity ($H_d$), and haplotype number ($H$) parameters

**Table 2**
Analysis of molecular variance (AMOVA) for components of haplotypic variation and paired $F_{ST}$ value for TMGMV main clades.

| Genomic Region | Source of variation | d.f. | Sum of squares | Variance components | Percentage of variation (%) | $F_{ST}$ |
|---|---|---|---|---|---|---|
| **Full genome** | Among clades | 2 | 739,056 | 43,88588 Va | 46,85 | 0,46850 |
| | Within clades | 25 | 1244,694 | 49,78778 Vb | 53,15 | |
| | Total | 27 | 1983,750 | 93,67366 | | |
| **p126** | Among clades | 2 | 458,722 | 27,38784 Va | 47,87 | 0,47873 |
| | Within clades | 25 | 745,528 | 29,82111 Vb | 52,13 | |
| | Total | 27 | 1204,250 | 57,20895 | | |
| **p183** | Among clades | 2 | 618,337 | 36,88127 Va | 47,69 | 0,47685 |
| | Within clades | 25 | 1011,556 | 40,46222 Vb | 52,31 | |
| | Total | 27 | 1629,893 | 77,34350 | | |
| **MP gene** | Among clades | 2 | 83,446 | 4,45666 Va | 41,75 | 0,41750 |
| | Within clades | 27 | 167,888 | 6,21806 Vb | 58,25 | |
| | Total | 29 | 251,333 | 10,67473 | | |
| **CP** | Among clades | 1 | 100,293 | 2,22162 Va | 42,22 | |
| | Within clades | 96 | 291,881 | 3,04042 Vb | 57,78 | 0,42220 |
| | Total | 97 | 392,173 | 5,26204 | | |

d.f.: degrees of freedom.Statistical significance: $P < 0.05$ Full-genome, p183, and p126 genes (n = 28; three main clades), MP gene (n = 30; three main clades), CP gene (n = 98, two main clades).

according to all variants and evolutionary clades that formed at the specific gene regions and complete genome level. The mean nucleotide diversities for all populations of the MP, p126, p183, and CP gene regions were 0,02263, 0,02678, 0,02474, and 0,01685, respectively. In addition, the main clade populations' nt diversity of TMGMV gene regions ranged from 0,00817 to 0,02803. Moreover, haplotype diversity ($H_d$) for gene regions varied between 0,947 and 1000. Therefore, it can be concluded that there was remarkable genetic variation among TMGMV variants due to low nt diversity and high haplotype diversity values (Table 3).

For each specific gene region and the entire genome, $d_N/d_S$ ($\omega$) ratios were calculated to determine whether positive or negative selection pressure influenced the structure of TMGMV populations. All $d_N/d_S$ values for each gene region were less than 1, and these values indicated the presence of negative selection pressure in the p126, p183, MP, and CP gene regions. More specifically, there was a stronger negative selection in the MP and CP gene regions ($\omega$ = 0,0818–0,2010) than in the p126 and p183 gene regions ($\omega$ = 0,1641–0,2627). On the other hand, the complete genome ω ratios indicated positive selection in all populations and clade I variants ($\omega$ = 1,0351–1,0455) and weak negative selection in clades II and III ($\omega$ = 0,9377–0,9576) (Table 3).

### 3.6. Neutrality tests and population differentiation

Neutrality tests were applied using Tajima's *D*, Fu and Li's D* & F* statistics to predict evolutionary changes operating on the complete genome and each gene region of TMGMV populations. The test results for both the MP and CP gene regions were statistically significant with negative values for MP-clade I/all variants ($P < *0.05$) and CP-clade I/II/all variants ($P < *0.05$ and $P < **0.02$) (Table 4). These findings might be interpreted as evidence of recent population expansion, genetic hitchhiking, or background selection. Besides, statistically non-significant negative and positive values were calculated for the p126 and p183 gene regions. Similarly, the results obtained for the whole genome had non-significant negative and positive values, except for Tajima's D test result for clade I (Table 4). These findings indicated that

**Table 3**
Summary of the genetic diversity analyses of TMGMV from main clade populations.

| Region | N | Pi | H | H*d* | $d_N/d_S$ |
|---|---|---|---|---|---|
| **Full genome** | | | | | |
| Clade I | 18 | 0,01505 | 18 | 1000 | 1,03511 |
| Clade II | 4 | 0,02469 | 4 | 1000 | 0,95761 |
| Clade III | 6 | 0,01235 | 6 | 1000 | 0,93773 |
| All | 28 | 0,02324 | 28 | 1000 | 1,04553 |
| **p126** | | | | | |
| Clade I | 18 | 0,01718 | 18 | 1000 | 0,1641 |
| Clade II | 4 | 0,02803 | 4 | 1000 | 0,1750 |
| Clade III | 6 | 0,01425 | 6 | 1000 | 0,2236 |
| All | 28 | 0,02678 | 28 | 1000 | 0,1641 |
| **p183** | | | | | |
| Clade I | 18 | 0,01588 | 18 | 1000 | 0,2015 |
| Clade II | 4 | 0,02609 | 4 | 1000 | 0,2232 |
| Clade III | 6 | 0,01291 | 6 | 1000 | 0,2627 |
| All | 28 | 0,02474 | 28 | 1000 | 0,2236 |
| **MP** | | | | | |
| Clade I | 19 | 0,01587 | 18 | 0,994 | 0,1346 |
| Clade II | 5 | 0,02179 | 5 | 1000 | 0,0818 |
| Clade III | 6 | 0,01332 | 6 | 1000 | 0,1060 |
| All | 30 | 0,02263 | 29 | 0,998 | 0,1590 |
| **CP** | | | | | |
| Clade I | 65 | 0,01492 | 48 | 0,985 | 0,09957 |
| Clade II | 33 | 0,00817 | 18 | 0,947 | 0,20104 |
| All | 98 | 0,01685 | 66 | 0,988 | 0,08865 |

**Table 4**
The neutrality test result of TMGMV from main clade populations.

| Region | Tajima's D | Fu and Li F* | Fu and Li D* |
|---|---|---|---|
| **Full genome** | | | |
| Clade I | −1,79967* | −2,06123 ns | −1,76058 ns |
| Clade II | −0,20887 ns | −0,20553 ns | −0,18667 ns |
| Clade III | 0,36608 ns | 0,55082 ns | 0,53746 ns |
| All | −1,54288 ns | −2,08247 ns | −1,87000 ns |
| **P126 Region** | | | |
| Clade I | −1,61435 ns | −1,76056 ns | −1,47173 ns |
| Clade II | −0,21470 ns | −0,19869 ns | −0,17809 ns |
| Clade III | 0,21765 ns | 0,38622 ns | 0,38782 ns |
| All | −1,43179 ns | −1,94414 ns | −1,75024 ns |
| **P183 Region** | | | |
| Clade I | −1,71803 ns | −1,92263 ns | −1,62580 ns |
| Clade II | −0,17861 ns | −0,16831 ns | −0,15132 ns |
| Clade III | 0,35162 ns | 0,53188 ns | 0,51969 ns |
| All | −1,47715 ns | −1,98563 ns | −1,78012 ns |
| **MP Region** | | | |
| Clade I | −2,11752* | −2,79542* | −2,52058* |
| Clade II | −0,40592 ns | −0,43846 ns | −0,40592 ns |
| Clade III | 0,12077 ns | 0,28083 ns | 0,29451 ns |
| All | −1,82890* | −2,71703** | −2,53127** |
| **CP Region** | | | |
| Clade I | −1,90712* | −2,96725* | −2,63956* |
| Clade II | −1,90712* | −2,75664* | −2,54432* |
| All | −2,12324* | −3,55514** | −3,58054 ** |

Statistical significance: *, $P < 0.05$; **, $P < 0.02$; **, $P < 0.01$.

TMGMV populations were going through bottlenecks or balancing selection according to the whole genome.

Ks*, Z*, $S_{nn}$, $F_{st}$, and Nm independent statistical tests based on permutation with *1000* replicates, were used to evaluate genetic differentiation and migration between/within major clades originating from phylogenetic trees reflecting evolutionary relationships. The genetic distinction values for the complete TMGMV genome and each of its gene regions demonstrated genetic differentiation between the major phylogroups, and these results were supported by a statistically significant $S_{nn}$ test result. Aside from MP clade I and II, the absolute $F_{st}$ value across the clades was higher than 0.33 and the absolute migration rate was lower than 1 (Table 5). In addition, these results were confirmed by a non-parametric permutation analysis based on *10,000* repetitions and the genetic differentiation between TMGMV phylogroups using the F-statistic. The values obtained from this test were ranked, with statistically significant values ranging from 0.41750 to 0.47865 (Table 5). All these findings were significantly supported by each independent statistical test, demonstrating the accuracy of the resulting evolutionary clades for the complete genome of TMGMV and for each gene region.

## 4. Discussion

Tobamovirus genetic variability studies have helped us better understand the genetic diversity of other viruses and the mechanisms that generate this diversity, as well as the factors that shape viral populations, host-virus-vector interactions, the evolutionary emergence of these interactions, and their host adaptations [54]. Therefore, genetic diversity studies are frequently used to comprehend the current population structures of many common and acute viruses and to forecast future population situations. Thus, in this study, the population genetic structure of TMGMV, which is the major agent in the largest tobacco cultivation areas in western Turkey (Fig. 1), was revealed by using genetic diversity parameters. In addition, the prevalence of tobamoviruses in the same region was determined using conventional molecular techniques.

Symptoms of infection caused by tobamoviruses can vary depending on the virus species, the host plant, and environmental conditions [55]. Typically, symptoms such as deformation, mosaics, and mottled patterns can be seen on the leaves of infected plants [56]. Additionally, the severity of the symptoms that tobamoviruses cause on the leaves or fruits of plants can range from mild to severe. However, it has been reported that TMGMV causes mild mosaic symptoms in *Nicotiana glauca* leaves, while TMV causes severe mosaic and leaf deformations [57]. In other studies, mild mottling was observed on *N. glauca* leaves in TMGMV infections, whereas leaf distortion and mosaic pattern symptoms were observed in TMV infections in tobacco [15,58]. While mild or severe mosaic symptoms were observed in the leaves of 300 tobacco samples in our study, the main symptoms consisted of mild mosaics (Fig. 2). According to the molecular test results, 75% of the samples were infected with TMGMV, and the symptoms progressed as a mild mosaic in tobacco, as previously reported in other studies [59]. Interestingly, in the molecular determination of ToBRFV, ToMMV, TMV, ToMV, and TMGMV agents, including important and common tobamovirus species, a single TMGMV infection was determined at a very high level in 225 of 300 plant samples. However, double infections of TMV + TMGMV in 21 plants and ToMV + TMGMV in only one plant were detected (Table 1). According to a previous study by Fraile et al. [57], it was reported that TMV colonized the *N. glauca* plant earlier or more quickly than TMGMV, and it also caused the second agent, TMGMV, to drop below the point at which harmful mutations were eliminated in the TMV population. According to our findings, the fact that TMGMV is more common than TMV in Turkey strongly supports the phenomenon called 'mutational melting' or 'Müller latch', which presumably causes TMV to disappear from its niche [57,60].

RNA-RNA recombination is one of the most effective forces shaping the genomes of plant RNA viruses and their genetic variability. Thus, it is now known that RNA recombination plays an important role in rearranging viral genes, repairing harmful mutations, and shaping the molecular evolutionary relationships of viral taxa [61]. Additionally, it has been reported that homologous recombination is effective in the evolution of tobamoviruses. More specifically, three recombinants containing each viral gene have been reported by He et al. [62], which may be derived from homologous or aberrant homologous recombination between TMV and ToMV. An analysis by the neighbor-joining method revealed inconsistent branching, supported by a moderate bootstrap relative (62%), to the middle of the TMGMV genome, namely the RdRp gene region [18]. However, this resulted in not providing conclusive evidence that TMGMV is derived from an intersubgroup recombinant [18]. In this study, the recombination analyses based on the nucleotide sequences of the Turkish and global variants have not produced any results that would support the existence of genetic recombination.

In previous studies, different techniques were used to examine the phylogenetic relationships of tobamoviruses, and today, owing to the rapid progress of sequencing technologies, molecular evolutionary analyses can be performed in detail based on nucleotide sequences. Furthermore, one of the most striking of these studies is the classification of tobamoviruses with a 126K protein comparison by Fraile and Garcia-Arenal [24]. More specifically, based on the CP gene region of TMGMV, it was reported that the populations clustered into a single group and did not reflect distribution according to host or geographical origin in the phylogenetic study conducted with 26 variants from Spain and 27 variants from different countries of the world [63]. Herein, molecular evolutionary analyses were performed using the nucleotide sequences of the entire TMGMV genome and the four significant protein-coding gene regions that comprise this genome. Thus, there were three distinct lineages that emerged based on the p126 and p183 gene regions of TMGMV: Clade I, II, and III. Furthermore, phylogenies of both gene regions showed striking similarities in the distributions of all viral

**Table 5**
Genetic differentiation and gene flow estimates for TMGMV main clades, based on the complete genome, CP, MP, p126, and p183genes.

| Genomic region | Comparisons | $K_s$* (*P* value) | *Z** (*P* value) | *Snn* (*P* value) | $F_{st}$ | Nm |
|---|---|---|---|---|---|---|
| **MP gene** | Clade I (n = 19)/Clade II (n = 5) | 2,50128 (0,0050 ***) | 4,37150 (0,0000 ***) | 0,97222 (0,0000 ***) | 0,32243 | 0,53 |
| | Clade I (n = 19)/Clade III (n = 6) | 2,39089 (0,0000 ***) | 4,24901 (0,0000 ***) | 1,00000 (0,0000 ***) | 0,48376 | 0,27 |
| | Clade II (n = 5)/Clade III (n = 6) | 2,45789 (0,0020 **) | 2,48637 (0,0020 **) | 1,00000 (0,0030 **) | 0,41074 | 0,36 |
| **p126 gene** | Clade I (n = 18)/Clade II (n = 4) | 4,03132 (0,0010 **) | 4,17713 (0,0010 **) | 1,00000 (0,0010 **) | 0,34521 | 0,47 |
| | Clade I (n = 18)/Clade III (n = 6) | 3,91288 (0,0000 ***) | 4,08450 (0,0000 ***) | 1,00000 (0,0000 ***) | 0,54801 | 0,21 |
| | Clade II (n = 4)/Clade III (n = 6) | 3,96629 (0,0020 **) | 2,25573 (0,0020 **) | 1,00000 (0,0170 *) | 0,41358 | 0,35 |
| **p183 gene** | Clade I (n = 18)/Clade II (n = 4) | 4,32043 (0,0010 **) | 4,16918 (0,0010 **) | 1,00000 (0,0010 **) | 0,34449 | 0,48 |
| | Clade I (n = 18)/Clade III (n = 6) | 4,18419 (0,0000***) | 4,07132 (0,0000***) | 1,00000 (0,0000 ***) | 0,55423 | 0,20 |
| | Clade II (n = 4)/Clade III (n = 6) | 4,21152 (0,0040 **) | 2,24148 (0,0040 **) | 1,00000 (0,0010*) | 0,41176 | 0,36 |
| **Full genome** | Clade I (n = 18)/Clade II (n = 4) | 4,52316 (0,0000 ***) | 4,17302 (0,0010*) | 1,00000 (0,0000 ***) | 0,33173 | 0,50 |
| | Clade I (n = 18)/Clade III (n = 6) | 4,38618 (0,0000 ***) | 4,07684 (0,0000 ***) | 1,00000 (0,0000 ***) | 0,54666 | 0,21 |
| | Clade II (n = 4)/Clade III (n = 6) | 4,41128 (0,0090**) | 2,23611 (0,0060**) | 1,00000 (0,0180 *) | 0,40939 | 0,36 |
| **CP gene** | Clade I (n = 65)/Clade II (n = 33) | 1,81395 (0,0000 ***) | 6,92473 (0,0000 ***) | 0,99405 (0,0000 ***) | 0,44637 | 0,31 |

PM test; Probability obtained by the permutation test with 1000 replicates)ns, not significant; *, 0.01 < P < 0.05; **, 0.001 < P < 0.01; ***, P < 0.001.

variants located in major lineages as well as tree topologies (Fig. 3). However, the main lineages of the phylogenetic tree that emerged for the complete genome showed significant similarity with the major clades formed in the p126 and p183 gene regions. Except for the clustering of variants from Germany and Italy, the molecular evolutionary tree of the MP gene region showed a significant match with the tree topologies of the p126 gene, the p183 gene, and the complete genome. The phylogenetic tree of the CP gene region constructed by Bera et al. [63] differed from that obtained in the present study. This difference may be explained by recent increases in the number of TMGMV variants uploaded to the GenBank, high nucleotide homology among these variants, and various methods and algorithms used in phylogenetic analyses without the use of an outgroup. In addition, the evolutionary tree derived from any gene region other than the CP gene region did not reflect any geographic or host-related clustering. Furthermore, with the exception of one variant from China, all members of clade II in the CP gene region were variants from Spain. Therefore, Clade II of the CP gene region was referred to as "the Spanish Clade" in this study. Consequently, phylogenetic analyses revealed that variants obtained from Western Turkey, with the exception of the CP gene region, clustered together with global variants in two main lineages (Clade I and III) at each gene and complete genome levels (Fig. 3).

Purifying (negative) selection increases the speed at which harmful gene mutations are removed and a stable population genetic structure is formed, whereas natural selection is another significant evolutionary mechanism and the driving force behind virus population variability [64]. Herein, $\omega$ ratio values demonstrated that four TMGMV genes, including MP, p126, p183, and CP, were under purifying selection. Additionally, positive selection for Spanish populations of TMGMV and negative selection for other populations have been reported by Fraile et al. [27]. Generally, it has been reported that the capsid proteins of vector-borne plant viruses are subject to more purifying selection for amino acid substitution than viruses transmitted by other routes [65]. Moreover, several studies have shown that vector-borne plant viruses are under significant purifying selection pressure, especially for the CP gene [66,67]. On the other hand, this study also showed that the CP gene of TMGMV, which is not a vector-borne virus, has a strong purifying selection pressure relative to its other gene regions. This situation may be associated with the adaptation of stable TMGMV particles to environmental and ecological changes. Furthermore, the negative selection found on the p126 and p183 genes, which are silencing suppressor proteins involved in viral genome replication, could be linked to the RNA interference mechanism of the host plant.

To assess the expansion status of the TMGMV population, haplotype diversity and nucleotide diversity were calculated between phylogroups and for all populations. The values obtained by genetic diversity analyses are detailed in Table 3, and the results showed extremely low nucleotide diversity ($\pi$) (0,00817–0,02803) and high haplotype diversity (Hd) (0,947–1000) in each specific gene region and the complete genome level (Table 3). These results may indicate the presence of a large number of closely related haplotypes in the population and that this population may have undergone a recent expansion. Moreover, the values we obtained in our study are similar to the average nucleotide diversity of each gene reported for other plant viruses such as Beet black scorch virus (genus: Betanecrovirus; BBSV) populations [68,69]. Thus, the population structure results obtained in our study are quite consistent, with the finding that TMGMV populations have high genetic stability, as previously reported [25].

Neutrality tests yielded negative results with statistically significant values in the MP and CP genes (Table 4). This finding could demonstrate that the recent expansion in TMGMV populations, as well as their positive selection associations, are linked to the functional and structural MP and CP gene regions. Furthermore, statistically non-significant negative and positive values were calculated for the p126, and p183 gene regions and the complete genome, indicating that TMGMV populations were going through bottlenecks or balancing selection. This situation can be attributed to the general notion that extreme bottlenecks were detected at the systemic leaf colonization stage in all investigated plant viruses and that some unknown barriers are an impediment to the development of all plant viruses [70]. Population bottlenecks in the host plant can cause losses of variation in its disease-resistance loci and have significant consequences on the ability of these populations to adapt to pathogen pressure [71]. The changes in the host metabolism caused by the p126 and p183 genes, which are responsible for the replication of the viral genome and have the silencing suppressor protein feature, strongly support the idea that the populations enter the bottleneck, which has been suggested in terms of both the plant and the viral agent. Thus, these reactions at the molecular level can both shape population structures and affect the severity of symptoms, mutually with the host.

One of the key issues in population genetics is differentiation [69]. In this research, independent test statistic values were calculated to support the distinctions between the main lineages revealed as a consequence of molecular evolutionary analyses conducted on the complete TMGMV genome and four specific gene regions. Ks*, Z*, Snn, FST, and Nm values were given in detail in Table 5. Values between the variance of allele frequencies ($F_{ST}$) and migration rates (Nm) supported differences between main lineages with infrequent gene flow results ($F_{ST}>$ 0.33 or Nm $<$ 1). In addition, other Ks*, Z*, and Snn test values statistically supported the clustering of variants in different main lineages (Table 5). On the other hand, an AMOVA test was applied to determine whether the main clusters formed in the phylogenetic trees for each gene region and the complete genome of TMGMV are a reflection of genetic differences and variations within or among major clades (Table 2). Thus, the percentages of molecular variance obtained from the phylogenetic groups of each gene and the whole genome evaluated in terms of genetic difference showed very parallel and close values with each other. Consequently, it was assumed that each gene region contributed in a consistent manner to the total genetic variation when all of these values were considered together. In addition, $F_{ST}$ test results based on the source of variations, including among and within main clades, also confirmed the lineages shown by phylogenetic analyses with statistical significance.

## 5. Conclusion

The proposed phylogenetic structure of tobamoviruses supports the idea that these viruses and their hosts have co-evolved since the hosts diverged from a common ancestor [72]. On the other hand, studies on the genetic diversity and population structure of TMGMV, one of the important members of this genus, date back almost 30 years. Thus, in this study, the population genetic structure of TMGMV was investigated in detail based on the nucleotide sequences of both naturally infected Turkish tobacco variants and global variants. Comprehensively, the molecular evolutionary relationships of TMGMV for both the complete genome and for each different gene region revealed the existence of three major lineages (Clade I/II/III) except for the CP gene (Clade I/II). Although all TMGMV variants have high nucleotide similarity, analyses of population genetic and molecular variance strongly support the existence of these main lineages. Moreover, variants from Turkey were clustered in two different lineages. We attribute this finding to the stability of the virus, the approximately 400-year history of tobacco cultivation in Anatolian areas, and the prospect that the virus may have persisted in plant residues. Consequently, although plant viruses have a high genetic variation potential, we have determined through comprehensive analyses that the purifying pressures on the structural and functional genes of TMGMV shape the population structure and that the genetic stability rule is effective in these populations. However, for future studies, more TMGMV variants representing different geographic regions will reveal more information about the population genetic structure and the effects of genes on this structure.

## Ethical approval

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

## Credit authorship contribution statement

All authors contributed to the writing of the manuscript. AK performed field and laboratory works. FRZ performed bioinformatic analyses, supervised the conceptualization, and wrote the first draft of MS. SK get funding and supervised the study. All authors have read and approved the final manuscript.

## Funding

This study was supported by a grant from the TUBITAK-TOVAG (project # 119O625).

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pmpp.2023.102008.

## References

[1] TURKSTATS, 2020. https://biruni.tuik.gov.tr/medas/?kn=92&locale=tr.
[2] İ. Özkul, Y. Sarı, Türkiye'de tütün sektörünün durumu, sorunları ve çözüm önerileri. 2, Ulusal İktisat Kongresi (2008) 1–22 (in Turkish).
[3] S. Daşdemir, A Research on Determination of Suitable Growing Sites for Planting Sone Tobacco Species Using Remote Sensing Technique and Querying of These Ina Geographical Information System Environment, Master Thesis, Ege University, in Turkish, 2006, 180 pp.
[4] A.M. King, E. Lefkowitz, M.J. Adams, E.B. Carstens (Eds.), Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses, vol. 9, Elsevier, 2011.
[5] M.J. Adams, S. Adkins, C. Bragard, D. Gilmer, D. Li, S.A. MacFarlane, et al., ICTV virus taxonomy profile: Virgaviridae, J. Gen. Virol. 98 (2017) 1999, https://doi.org/10.1099/jgv.0.000884.
[6] S. Adkins, I. Kamenova, D. Achor, D.J. Lewandowski, Biological and molecular characterization of a novel tobamovirus with a unique host range, Plant Dis. 87 (2003) 1190–1196, https://doi.org/10.1094/PDIS.2003.87.10.1190.
[7] S. Adkins, I. Kamenova, E.N. Rosskopf, D.J. Lewandowski, Identification and characterization of a novel tobamovirus from tropical soda apple in Florida, Plant Dis. 91 (2007) 287–293, https://doi.org/10.1094/PDIS-91-3-0287.
[8] N. Ilmberger, P. Willingmann, G. Adam, C. Heinze, A subgroup 1 Tobamovirus isolated from Brugmansia sp. and its detection by RT-PCR, J. Phytopathol. 155 (2007) 326–332, https://doi.org/10.1111/j.1439-0434.2007.01235.x.
[9] Y. Li, Y. Wang, J. Hu, et al., The complete genome sequence, occurrence and host range of Tomato mottle mosaic virus Chinese isolate, Virol. J. 14 (2017) 15, https://doi.org/10.1186/s12985-016-0676-2.
[10] B.E. Min, B.N. Chung, M.J. Kim, J.H. Ha, B.Y. Lee, K.H. Ryu, Cactus mild mottle virus is a new cactus-infecting tobamovirus, Arch. Virol. 151 (2006) 13–21, https://doi.org/10.1007/s00705-005-0617-7.
[11] Y.S. Song, B.E. Min, J.S. Hong, M.J. Rhie, M.J. Kim, K.H. Ryu, Molecular evidence supporting the confirmation of maracuja mosaic virus as a species of the genus tobamovirus and production of an infectious cDNA transcript, Arch. Virol. 151 (2006) 2337–2348, https://doi.org/10.1007/s00705-006-0823-y.
[12] A. Gibbs, Evolution and origins of tobamoviruses, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 354 (1383) (1999) 593–602.
[13] M.I. Font, M.C. Córdoba-Sellés, M.C. Cebrián, J.A. Herrera-Vásquez, A. Alfaro-Fernández, A. Boubaker, I. Soltani, C. Jordá, First report of tobacco mild green mosaic virus Infecting Capsicum annuum in Tunisia, Plant Dis. 93 (7) (2009) 761, https://doi.org/10.1094/PDIS-93-7-0761B.
[14] A. Skelton, C. Sansford, Pest Risk Analysis for Tobacco Mild Green Mosaic Virus, The Food and Environmental Research Agency, 2012. https://secure.fera.defra.gov.uk/phiw/riskRegister/downloadExternalPra.cfm?id=3924. (Accessed 6 March 2019).
[15] G.M. Favara, V.M. Camelo-García, T. Nagata, et al., Tobacco mild green mosaic virus found naturally infecting Nicotiana glauca in Brazil, Australas. Plant Dis. Notes 14 (2019) 13, https://doi.org/10.1007/s13314-019-0345-6.
[16] S.C. Balsak, S.G. Avcu, N. Buzkan, Detection and molecular characterization of tobacco mild green mosaic virus isolates infecting peppers in Turkey, J. Agricult. Nat. 25 (2) (2022) 490–496.
[17] A. Karanfil, M. Sarı, S. Korkmaz, First report of tobacco mild green mosaic virus in Turkey, J. Plant Pathol. 102 (2020) 547, https://doi.org/10.1007/s42161-019-00436-x.
[18] R.T. Lartey, T.C. Voss, U. Melcher, Tobamovirus evolution: gene overlaps, recombination, and taxonomic implications, Mol. Biol. Evol. 13 (10) (1996) 1327–1338, https://doi.org/10.1093/oxfordjournals.molbev.a025579.
[19] L.Y. Wang, S.S. Lin, T.H. Hung, T.K. Li, N.C. Lin, T.L. Shen, Multiple domains of the tobacco mosaic virus p126 protein can independently suppress local and systemic RNA silencing, Mol. Plant Microbe Interact. 25 (5) (2012) 648–657, https://doi.org/10.1094/MPMI-06-11-0155.
[20] D. Silhavy, J. Burgyán, Effects and side-effects of viral RNA silencing suppressors on short RNAs, Trends Plant Sci. 9 (2) (2004) 76–83, https://doi.org/10.1016/j.tplants.2003.12.010.
[21] A. Niehl, M. Heinlein, Cellular pathways for viral transport through plasmodesmata, Protoplasma 248 (2011) 75–99, https://doi.org/10.1007/s00709-010-0246-1.
[22] B. Moury, C. Desbiez, M. Jacquemond, H. Lecoq, Genetic diversity of plant virus populations: towards hypothesis testing in molecular epidemiology, Adv. Virus Res. 67 (2006) 49–87, https://doi.org/10.1016/S0065-3527(06)67002-4.
[23] M. Adhab, C. Angel, A. Rodriguez, M. Fereidouni, L. Király, K. Scheets, J. E. Schoelz, Tracing the lineage of two traits associated with the coat protein of the Tombusviridae: silencing suppression and HR elicitation in Nicotiana species, Viruses 11 (7) (2019) 588.
[24] A. Fraile, F.A. García-Arenal, Classification of the tobamoviruses based on comparisons among their 126K proteins, J. Gen. Virol. 71 (1990) 2223–2228, https://doi.org/10.1099/0022-1317-71-10-2223.
[25] E. Rodríguez-Cerezo, S.F. Elena, A. Moya, et al., High genetic stability in natural populations of the plant RNA virus tobacco mild green mosaic virus, J. Mol. Evol. 32 (1991) 328–332, https://doi.org/10.1007/BF02102191.
[26] A. Moya, E. Rodriguez-Cerezo, F. García-Arenal, Genetic structure of natural populations of the plant RNA virus tobacco mild green mosaic virus, Mol. Biol. Evol. 10 (1993), https://doi.org/10.1093/oxfordjournals.molbev.a040019, 449-449.
[27] A. Fraile, J.M. Malpica, M.A. Aranda, E. Rodríguez-Cerezo, F. García-Arenal, Genetic diversity in tobacco mild green mosaic tobamovirus infecting the wild plant Nicotiana glauca, Virology 223 (1) (1996) 148–155, https://doi.org/10.1006/viro.1996.0463.
[28] R. Li, R. Mock, Q. Huang, J. Abad, J. Hartung, G. Kinard, A reliable and inexpensive method of nucleic acid extraction for the PCR-based detection of diverse plant pathogens, J. Virol. Methods 154 (1–2) (2008) 48–55, https://doi.org/10.1016/j.jviromet.2008.09.008.
[29] B.M. Muhire, A. Varsani, D.P. Martin, SDT: a virus classification tool based on pairwise sequence alignment and identity calculation, PLoS One 9 (2014), 0108277, https://doi.org/10.1371/journal.pone.0108277.
[30] K. Tamura, G. Stecher, S. Kumar, MEGA 11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (7) (2021) 3022–3027, https://doi.org/10.1093/molbev/msab120.
[31] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680, https://doi.org/10.1093/nar/22.22.4673.
[32] M. Nei, S. Kumar, Molecular Evolution and Phylogenetics, Oxford University Press, New York, 2000.
[33] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425, https://doi.org/10.1093/oxfordjournals.molbev.a040454.
[34] A. Rzhetsky, M. Nei, A simple method for estimating and testing minimum evolution trees, Mol. Biol. Evol. 9 (1992) 945–967, https://doi.org/10.1093/oxfordjournals.molbev.a040771.
[35] K. Tamura, Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases, Mol. Biol. Evol. 9 (1992) 678–687, https://doi.org/10.1093/oxfordjournals.molbev.a040752.
[36] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791.
[37] D. Martin, E. Rybicki, RDP: detection of recombination amongst aligned sequences, Bioinformatics 16 (6) (2000) 562–563, https://doi.org/10.1093/bioinformatics/16.6.562.
[38] M.O. Salminen, J.K. Carr, D.S. Burke, F.E. Mccutchan, Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning, AIDS Res. Hum. Retrovir. 11 (11) (1995) 1423–1425, https://doi.org/10.1089/aid.1995.11.1423.
[39] S.A. Sawyer, GENECONV: A Computer Package for the Statistical Detection of Gene Conversion, Washigton University, St Louis, USA, 1999. Distributed by the author. Department of Mathematics. Available at: http://www.math.wustl.edu/sawyer.
[40] J.M. Smith, Analyzing the mosaic structure of genes, J. Mol. Evol. 34 (1992) 126–129.
[41] D. Posada, K.A. Crandall, Evaluation of methods for detecting recombination from DNA sequences: computer simulations, Proc. Natl. Acad. Sci. USA 98 (24) (2001) 13757–13762, https://doi.org/10.1073/pnas.241370698.

[42] M.J. Gibbs, J.S. Armstrong, A.J. Gibbs, Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences, Bioinformatics 16 (7) (2000) 573–582, https://doi.org/10.1093/bioinformatics/16.7.573.
[43] D.P. Martin, A. Varsani, P. Roumagnac, G. Botha, S. Maslamoney, T. Schwab, et al., RDP5: a computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets, Virus Evol. 7 (1) (2021) veaa087.
[44] L. Excoffier, P. Smouse, J. Quattro, Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data, Genetics 131 (1992) 479–491, https://doi.org/10.1093/genetics/131.2.479.
[45] L. Excoffier, H.E.L. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour. 10 (2010) 564–567, https://doi.org/10.1111/j.1755-0998.2010.02847.x.
[46] B.S. Weir, C.C. Cockerham, Estimating F-statistics for the analysis of population structure, Evolution 38 (1984) 1358–1370, https://doi.org/10.1111/j.1558-5646.1984.tb05657.x.
[47] J. Rozas, A. Ferrer-Mata, J.C. Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S. E. Ramos-Onsins, A. Sánchez-Gracia, DnaSP 6: DNA sequence polymorphism analysis of large data sets, Mol. Biol. Evol. 34 (12) (2017) 3299–3302, https://doi.org/10.1093/molbev/msx248.
[48] M. Nei, Molecular Evolutionary Genetics, Columbia University Press, New York, 1987.
[49] F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, Genetics 123 (3) (1989) 585–595, https://doi.org/10.1093/genetics/123.3.585.
[50] Y.X. Fu, W.H. Li, Statistical tests of neutrality of mutations, Genetics 133 (3) (1993) 693–709, https://doi.org/10.1093/genetics/133.3.693.
[51] R.R. Hudson, A new statistic for detecting genetic differentiation, Genetics 155 (2000) 2011–2014, https://doi.org/10.1093/genetics/155.4.2011.
[52] R.R. Hudson, D.D. Boos, N.L. Kaplan, A statistical test for detecting geographic subdivision, Mol. Biol. Evol. 9 (1992) 138–151, https://doi.org/10.1093/oxfordjournals.molbev.a040703.
[53] K. Ohshima, R. Nomiyama, S. Mitoma, Y. Honda, R. Yasaka, K. Tomimura, Evolutionary rates and genetic diversities of mixed potyviruses in narcissus, Infect. Genet. Evol. 45 (2016) 213–223, https://doi.org/10.1016/j.meegid.2016.08.036.
[54] A. Fraile, F. García-Arenal, Tobamoviruses as models for the study of virus evolution, Adv. Virus Res. 102 (2018) 89–117, https://doi.org/10.1016/bs.aivir.2018.06.006.
[55] R. Ilyas, M.J. Rohde, K.R. Richert-Pöggeler, H. Ziebell, To Be seen or not to Be seen: latent infection by tobamoviruses, Plants 11 (16) (2022) 2166, https://doi.org/10.3390/plants11162166.
[56] U. Melcher, D.J. Lewandowski, W.O. Dawson, Tobamoviruses (Virgaviridae), Elsevier, Amsterdam, The Netherlands, 2021, ISBN 9780123744104, 2021.
[57] A. Fraile, F. Escriu, M.A. Aranda, J.M. Malpica, A.J. Gibbs, F. García-Arenal, A century of tobamovirus evolution in an Australian population of Nicotiana glauca, J. Virol. 71 (11) (1997) 8316–8320, https://doi.org/10.1128/JVI.71.11.8316-8320.1997.
[58] A. Gunay, M. Usta, First investigation of five tobacco viruses using pcr based methods in tobacco plants grown in Adiyaman, Turkey, Fresenius Environ. Bull. 29 (12) (2020) 11624–11632.
[59] H.H. McKinney, Mosaic diseases in the canary islands, west africa, and Gibraltar, J. Agric. Res. 39 (1929) 557.
[60] M. Lynch, R. Bürger, D. Butcher, W. Gabriel, The mutational meltdown in asexual populations, J. Hered. 84 (5) (1993) 339–344, https://doi.org/10.1093/oxfordjournals.jhered.a111354.
[61] J. Sztuba-Solińska, A. Urbanowicz, M. Figlerowicz, J.J. Bujarski, RNA-RNA recombination in plant virus replication and evolution, Annu. Rev. Phytopathol. 49 (2011) 415–443, https://doi.org/10.1146/annurev-phyto-072910-095351.
[62] M. He, C.Q. He, N.Z. Ding, Natural recombination between tobacco and tomato mosaic viruses, Virus Res. 163 (1) (2011) 374–379, https://doi.org/10.1016/j.virusres.2011.09.006.
[63] S. Bera, A. Fraile, F. García-Arenal, Analysis of fitness trade-offs in the host range expansion of an RNA virus, tobacco mild green mosaic virus, Virol. J. 92 (24) (2018), https://doi.org/10.1128/jvi.01268-18 e01268-18.
[64] M. Perez-Losada, M. Porter, K.A. Crandall, Methods for analyzing viral evolution, in: Plant Virus Evolution. M. J. Roossinck, Springer, Heidelberg, Berlin, Germany, 2008, pp. 165–204.
[65] E.R. Chare, E.C. Holmes, Selection pressures in the capsid genes of plant RNA viruses reflect mode of transmission, J. Gen. Virol. 85 (2004) 3149–3157, https://doi.org/10.1099/vir.0.80134-0.
[66] F. Randa-Zelyüt, F. Ertunç, Population genetic analysis of lettuce big-vein disease viruses and their vector fungi Olpidium virulentus in Ankara province, Turkey, Physiol. Mol. Plant Pathol. 113 (2021), 101593, https://doi.org/10.1016/j.pmpp.2020.101593.
[67] S.R. Sun, J.S. Chen, E.Q. He, M.T. Huang, H.Y. Fu, J.J. Lu, S.J. Gao, Genetic variability and molecular evolution of maize yellow mosaic virus populations from different geographic origins, Plant Dis. 105 (4) (2021) 896–903, https://doi.org/10.1094/PDIS-05-20-1013-RE.
[68] F. Garcia-Arenal, A. Fraile, J.M. Malpica, Variability and genetic structure of plant virus populations, Annu. Rev. Phytopathol. 39 (2001) 157–186, https://doi.org/10.1146/annurev.phyto.39.1.
[69] S. Farzadfar, R. Pourrahim, Positive selection and recombination shaped the large genetic differentiation of Beet black scorch virus population, PLoS One 14 (4) (2019), e0215574, https://doi.org/10.1371/journal.pone.0215574.
[70] B. Monsion, R. Froissart, Y. Michalakis, S. Blanc, Large bottleneck size in Cauliflower Mosaic Virus populations during host plant colonization, PLoS Pathog. 4 (10) (2008), e1000174, https://doi.org/10.1371/journal.ppat.1000174.
[71] G. Gos, T. Slotte, S.I. Wright, Signatures of balancing selection are maintained at disease resistance loci following mating system evolution and a population bottleneck in the genus Capsella, BMC Evol. Biol. 12 (2012) 152, https://doi.org/10.1186/1471-2148-12-152.
[72] A.H. Stobbe, U. Melcher, M.W. Palmer, M.J. Roossinck, G. Shen, Co-divergence and host-switching in the evolution of tobamoviruses, J. Gen. Virol. 93 (2) (2012) 408–418.