



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

Zohreh Moradi ^{1*} and Mohsen Mehrvar²

¹Department of Plant Pathology, Faculty of Crop Sciences, Sari Agricultural Sciences and Natural Resources University, P.O. Box 578, Sari, Iran

²Department of Plant Pathology, Faculty of Agriculture, Ferdowsi University of Mashhad, P.O. Box 91779-1163, Mashhad, Iran

(Received on October 13, 2021; Revised on October 30, 2021; Accepted on November 2, 2021)

Alfalfa mosaic virus (AMV), an economically important pathogen, is present worldwide with a very wide host range. This work reports for the first time the infection of *Vinca minor* and *Wisteria sinensis* with AMV using RNA sequencing and reverse transcription polymerase chain reaction confirmation. *De novo* assembly and annotating of contigs revealed that RNA1, RNA2, and RNA3 genomic fragments consist of 3,690, 2,636, and 2,057 nucleotides (nt) for IR-VM and 3,690, 2,594, and 2,057 nt for IR-WS. RNA1 and RNA3 segments of IR-VM and IR-WS closely resembled those of the Chinese isolate HZ, with 99.23-99.26% and 98.04-98.09% nt identity, respectively. Their RNA2 resembled that of Canadian isolate CaM and American isolate OH-2-2017, with 97.96-98.07% nt identity. The P2 gene revealed more nucleotide diversity compared with other genes. Genes in the AMV genome were under dominant negative selection during evolution, and the P1 and coat protein (CP) proteins were subject to the strongest and weakest purifying selection, respectively. In the popula-

tion genetic analysis based on the CP gene sequences, all 107 AMV isolates fell into two main clades (A, B) and isolates of clade A were further divided into three groups with significant subpopulation differentiation. The results indicated moderate genetic variation within and no clear geographic or genetic structure between the studied populations, implying moderate gene flow can play an important role in differentiation and distribution of genetic diversity among populations. Several factors have shaped the genetic structure and diversity of AMV: selection, recombination/reassortment, gene flow, and random processes such as founder effects.

Keywords : Alfamovirus, evolutionary analysis, lesser periwinkle, RNA-seq, wisteria

*Corresponding author.

Phone) +98-(0)11-33687567, FAX) +98-(0)11-33687567

E-mail) z.moradi@sanru.ac.ir

ORCID

Zohreh Moradi

https://orcid.org/0000-0002-8989-6450

Handling Editor : Rae-Dong Jeong

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Articles can be freely viewed online at www.ppjonline.org.

Alfalfa mosaic virus (AMV), the type member of the genus *Alfamovirus* (family *Bromoviridae*) (Bujarski et al., 2012), has a tripartite genome composed of single-stranded positive-sense RNAs, of which RNA1 and RNA2 encode the viral replicase proteins P1 and P2, respectively, and RNA3 encode the movement protein (MP) and coat protein (CP) (Bol, 1999; van Dun et al., 1987). The CP is translated from a subgenomic messenger RNA4, which is synthesized during the replication of RNA3 (Smit and Jaspars, 1982). AMV can be transmitted by several aphid species in a non-persistent manner, and also by mechanical inoculation, seed and pollen of some plants (Edwardson and Christie, 1997; He et al., 2010; Hiruki and Hampton, 1990). AMV was first isolated from alfalfa (*Medicago sativa*) in the United States in 1931 (Weimer, 1931) and to date, it has occurred in several countries (including Australia, New

Zealand, Saudi Arabia, North and South America, France, England, Italy, Greece, Egypt, China, and Iran) with the various degree of economic losses (Al-Shahwan, 2002; Che et al., 2020; Fletcher, 2001; Maina et al., 2019; Massumi et al., 2012; Sawalha and Mansour, 1996). This virus can naturally infect 698 species of 167 genera in 71 families (Edwardson and Christie 1997; Fletcher, 2001; Jasper and Bos, 1980; Xu and Nie, 2006). In Iran, the incidence of AMV has been reported in more than 25 plant species in different regions by serological and in some cases molecular methods (Esfandiari et al., 2005; Golnaraghi et al., 2004; Hamzeh et al., 2010; Mangeli et al., 2019; Massumi et al., 2012; Pourrahim and Farzadfar 2015; Zenaddini et al., 2004). Although AMV has long been known as an important pathogen in Iran, it has not been well characterized at the entire genomic sequence, and the available sequence data are predominantly confined to CP sequence. *Vinca minor* (family *Apocynaceae*) commonly known as lesser or dwarf periwinkle, is a perennial subshrub (United States Department of Agriculture, 2020) extensively grown as a flowering evergreen ornamental and ground cover plant. Lesser Periwinkle is not only an ornamental plant with lilac-blue flowers but also a medicinal value producing important indole alkaloids especially vincamine, as the major alkaloid found in the leaves with cerebrovasodilatory and neuroprotective activity (Farahanikia et al., 2011; Vas and Gulyás, 2005). *Wisteria sinensis* (family *Fabaceae*), is a perennial flowering climbing species with extreme longevity (Jiang et al., 2011), which is used in Iran and several other countries as an ornamental plant. Up to now, no information for infection of lesser periwinkle and wisteria with AMV has been reported worldwide. High-throughput sequencing can be used to better understand plant diseases, especially when the viral etiology is unknown (Prabha et al., 2013). Deep sequencing of the transcriptome (RNA-seq) allows us to find new viruses, investigate their genome, and determine their single nucleotide polymorphisms (SNPs) within a replicating population (Adams et al., 2009; Wu et al., 2015). Studies concerning the genetic variability and molecular evolutionary history of AMV will be helpful to understand the effects of variation caused by mutation, recombination, selection pressure, and adaptation in viral populations which leads to developing breeding programmes and sustainable management schemes (Moradi and Mehrvar, 2019). Very little is known about the genetic structure and evolution of AMV in the world. Therefore, the general objectives of this study were to (1) determining the complete genome of AMV isolates obtained from lesser periwinkle and wisteria as its new hosts (2) the phylogenetic relationships of new isolates to those already

described elsewhere (3) comparative studies of the population genetic structure of AMV and determine the main evolutionary and demographic mechanisms responsible for the observed population genetic structure. The CP gene is one of the most common molecular markers for the investigation of genetic diversity and molecular evolution in many plant viruses (Cuevas et al., 2012; Gao et al., 2016; Song et al., 2019) and was therefore selected for this study.

Materials and Methods

RNA isolation, cDNA library preparation, RNA-sequencing, and bioinformatics analysis. In July to June 2020, *V. minor* and *W. sinensis* leaf samples exhibiting severe virus-like disease symptoms of yellow halo, bright yellow mottle, or mosaic (Fig. 1) were collected from urban areas Mashhad, Razavi Khorasan province, in northeast Iran. To identify the causal agent(s) of this disease, total RNA was extracted from the one symptomatic leaf sample of each plant using SV Total RNA Isolation Kit (Promega, Madison, WI, USA) according to the manufacturer's specifications. RNA-seq libraries were prepared by Illumina TruSeq Stranded Total RNA sample preparation kit with Ribo-Zero Plant rRNA Removal and sequenced using an Illumina NovaSeq 6000 (Macrogen, Seoul, Korea) with 2×151 bp paired-end reads. After adapter trimming and

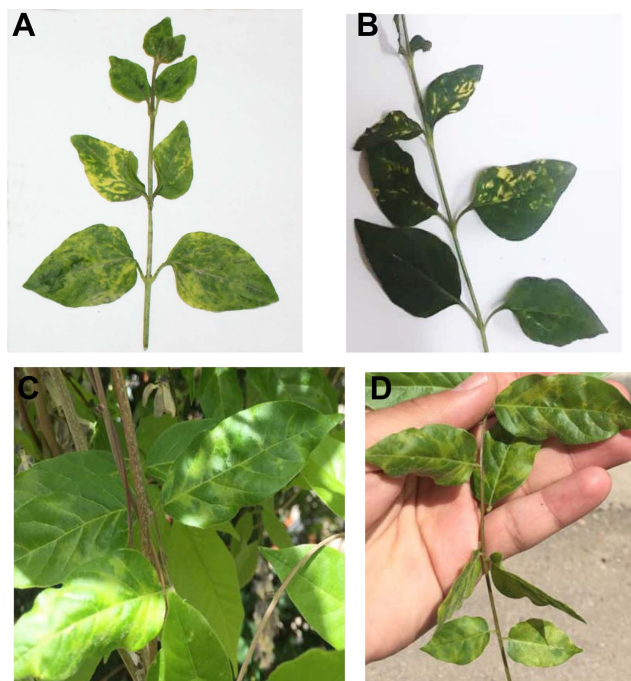


Fig. 1. Bright yellow mottle or mosaic symptoms of Alfalfa mosaic virus on *Vinca minor* (A, B) and *Wisteria sinensis* (C, D) leaves.

quality control, the clean reads of symptomatic plants were mapped to reads from healthy-looking plants. Unmapped reads were used for viral genome *de novo* assembly using CLC Genomics Workbench v.20 (CLC Bio, Qiagen, Aarhus, Denmark). All the contigs were subjected to BlastN and BlastX interrogation of the NCBI nucleotide and amino acid databases. Contigs that matched plant viruses were identified and imported to Geneious Prime v. 2019.1.3 (Biomatters, Auckland, New Zealand), and multiple alignments with reference sequences were performed using ClustalW (Thompson et al., 1994). The open reading frame prediction, annotation of the final sequences, gene translation, and prediction of deduced proteins were done in Geneious Prime v. 2019.1.3 (Biomatters). Motif searches were conducted in PROSITE (<http://www.expasy.ch/>), Pfam (<http://pfam.sanger.ac.uk/>), and CDD (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) databases. Reverse transcription polymerase chain reaction (RT-PCR) and Sanger sequencing method were used to further validate the occurrence of RNA-seq-detected virus in the samples with specific primers which were designed based on the CP region of the obtained contigs. Additional primers were designed from each viral genome sequence using the NCBI tool Primer-Blast (data not shown). SNPs were also determined using Geneious Prime.

Recombination/reassortment analysis. Assessment of the recombination/reassortment events in the nearly complete genomes of AMV from this study and those retrieved from GenBank ($n = 13$) was carried out using different algorithms implemented in RDP v.4.16 (Martin et al., 2015). A Bonferroni-corrected P -value cutoff of 0.05 and the option of “Sequences are linear” were selected. To reduce the possibility of obtaining false positives, recombination events supported by at least four methods with an associated P -value less than 10^{-6} were considered to be significant.

Sequence comparison and phylogenetic analysis. The nucleic acid sequences of P1, P2, MP, and CP genes of AMV isolates from Iran were aligned with the homologous sequences of other AMV isolates from the GenBank database by ClustalW using the IUB DNA weight matrix, whereas amino acid sequences were aligned by MUSCLE implemented in Geneious Prime version 2019.1.3 (Biomatters). Comparison of nucleotide and amino acid sequence identities were carried out using the Geneious Prime v. 2019.1.3 (Biomatters). The relationships of the four aligned open reading frame sequences were assessed using neighbor-joining (NJ) and maximum likelihood methods in MEGAX (Kumar et al., 2018). Branch support was evalu-

ated by Kimura’s two-parameter option, which was used to calculate 1,000 bootstrap replications. Genetic distances were estimated using the Kimura-2 parameter model in MEGAX and standard deviations were calculated by bootstrapping with 1,000 replicates.

Analysis of the extent of selection pressure on each cistron and across the sites. The nucleotide diversity (π) value and the extent of selective pressure were measured by comparing rates of synonymous (dS) and nonsynonymous (dN) substitutions among protein coding sequences using the DnaSP6 (Rozas et al., 2017). We further identified sites under positive selection using single-likelihood ancestor counting (SLAC), fast unbiased Bayesian approximation (FUBAR), fixed effects likelihood (FEL), internal fixed effects likelihood (IFEL), and mixed effects model of evolution (MEME) in Datamonkey (<http://www.datamonkey.org/>) (Pond and Frost, 2005).

Genetic diversity and population genetic differentiation based on CP sequences. The CP gene is an ideal target for virus diversity analysis due to its pivotal role in the virus life cycle, and there are large sequence datasets available in GenBank. Accordingly, CP gene sequences of 107 AMV isolates (three from this study and 104 retrieved from GenBank) (Supplementary Table 1) were used to determine phylogenetic correlation, to estimate genetic diversity parameters, to analyze subpopulation differentiation, and to investigate the important evolutionary factors shaping the genetic structure of AMV on a global scale. A codon-based alignment was performed using the MUSCLE algorithm (Edgar, 2004) included in MEGAX. Population genetic parameters were calculated using DnaSP version 6.10.04 software (Rozas et al., 2017) based on variant groups and geographical distributions. Tajima’s D (Tajima, 1989), Fu and Li’s D&F (Fu and Li, 1993) statistical tests, implemented in DnaSP6, were applied to evaluate the hypothesis of neutral selection acting on the CP gene. Genetic differentiation between populations was examined using several statistics: Ks^* , Kst^* , Z^* , and Snn based on permutation statistical tests with 1,000 replicates (Hudson, 2000). The null hypothesis of no genetic differentiation is rejected, if the test statistics (Ks^* , Kst^* , Z^* , and Snn) strongly supported by P -values < 0.05 . The degree of genetic differentiation or the level of gene flow between AMV populations was measured by the fixation index (F_{ST}) (Wright, 1951). F_{ST} can take values from 0, no genetic differentiation and complete gene flow, to 1, complete genetic differentiation as a consequence of null gene flow. Generally, $F_{ST} > 0.33$ implies infrequent gene flow, while $F_{ST} < 0.33$ indicates

frequent gene flow (Rozas et al., 2017).

Results

Molecular genomic characterization of AMV isolates.

Illumina sequencer produced 68,036,950 reads for a total of 10,118,891,750 bp in infected wisteria plant, and 62,554,812 reads for a total of 8,800,290,022 bp in infected vinca plant. *De novo* assembly generated 1,900 contigs with 112,390 to 3,511,021 reads (from wisteria) and 902 contigs with 9,281 to 2,734,240 reads (from vinca) mapped to the contigs of interest. For each surveyed plant, three contigs revealed AMV RNA genome segments (RNA1, RNA2, and RNA3) with GC contents of 42.70%, 41.70% (IR-WS, 41.80%), and 43.30%, respectively. RT-PCR using specific primers CP-F (5'-CATTGATCGGTA-ATGGGCCGT-3') and CP-R (5'-ATCCACCCAGTG-GAGGTCAGCA-3') (Massumi et al., 2012), which target CP gene, confirmed AMV in two samples. The AMV sequences of IR-VM RNA1 (3,690 nucleotides [nt]), RNA2 (2,636 nt), and RNA3 (2,057 nt) were deposited in the GenBank under the accession numbers MW014929, MW014930, and MW014931, respectively. Likewise, the AMV sequences of IR-WS RNA1 (3,690 nt), RNA2 (2,594 nt), and RNA3 (2,057 nt) were deposited in the GenBank under accession numbers MW014932, MW014933, and MW014934, respectively. In addition, RNA3 of another

isolate of AMV (IR-WS2) generated from wisteria and was deposited in GenBank with the accession numbers MW014935. RNA1 encoded the P1 gene, RNA2 encoded the P2 gene, and RNA3 encoded the CP and MP genes, which is typical of the AMV genome (Maina et al., 2019). Compared with other isolates, IR-VM and IR-WS contained an additional 19 nt in the 5'-untranslated region (5'-UTR). Interestingly, RNA1 of both isolates contained a short insertion of 47 nucleotides in the 3'-untranslated region (3'-UTR). In addition, the 3'-UTR of RNA2 sequence of IR-VM was divergent to length (contains 42 additional nt).

Sequence analysis. Pairwise comparisons showed that RNA1, RNA2, and RNA3 of IR-VM shared respectively 99.97%, 99.81%, and 99.95% nucleotide sequence identities with IR-WS. AMV-IR-VM and IR-WS shared nucleotide identities of 96.21-99.26%, 95.26-98.07%, and 94.26-98.09% with the 13 other AMV isolates obtained from GenBank in RNA1, RNA2, and RNA3 segments, respectively. AMV-IR-VM and IR-WS RNA1 shared the highest nucleotide identity with isolate HZ (HQ316635, China) (99.26% and 99.23%, respectively). Likewise, IR-VM and IR-WS RNA2 shared the highest nucleotide identity with isolates OH-2-2017 (MT669387, USA) (98.01%) and CaM (MK607975, Canada) (97.96% and 98.07%, respectively). RNA3 of IR-VM and IR-WS shared the

Table 1. Percent nucleotide sequence (and amino acid sequence) identities of IR-VM and IR-WS with other AMV isolates at the individual cistron level

Gene	IR-VM	IR-WS	HZ	Man-fredi	295	AU-SA80	Gyn	Lst	CaM	Ca175-1	175	Mint	Tec1	FERA 160224	OH-2-2017	
P1	IR-VM	100 (100)	99.97 (99.91)	99.41 (99.73)	99.14 (99.73)	99.11 (99.82)	98.91 (99.73)	98.85 (99.64)	98.91 (98.76)	98.52 (99.56)	98.08 (99.38)	98.11 (99.38)	97.66 (99.29)	96.45 (98.4)	96.72 (98.31)	97.75 (99.11)
	IR-WS	99.97 (99.91)	100 (100)	99.38 (99.64)	99.11 (99.64)	99.08 (99.73)	98.88 (99.64)	98.82 (99.56)	98.88 (98.67)	98.49 (99.47)	98.05 (99.29)	98.08 (99.29)	97.63 (99.2)	96.42 (98.31)	96.69 (98.22)	97.72 (99.02)
P2	IR-VM	100 (100)	100 (100)	95.19 (94.05)	97.81 (98.48)	95.78 (96.58)	95.57 (94.56)	96.88 (97.60)	97.81 (98.86)	98.02 (98.10)	95.27 (94.18)	95.32 (94.05)	97.89 (98.23)	96.08 (97.34)	97.81 (98.23)	97.97 (97.97)
	IR-WS	100 (100)	100 (100)	95.19 (94.05)	97.81 (98.48)	95.78 (96.58)	95.57 (94.56)	96.88 (97.60)	97.81 (98.86)	98.02 (98.10)	95.27 (94.18)	95.32 (94.05)	97.89 (98.23)	96.08 (97.34)	97.81 (98.23)	97.97 (97.97)
MP	IR-VM	100 (100)	100 (100)	98.78 (99.33)	97.67 (99.00)	97.34 (99.00)	97.45 (98.67)	97.45 (99.00)	97.56 (97.67)	97.79 (98.00)	97.12 (98.00)	97.12 (98.00)	97.67 (97.00)	94.13 (96.33)	93.58 (95.67)	98.01 (98.00)
	IR-WS	100 (100)	100 (100)	98.78 (99.33)	97.67 (99.00)	97.34 (99.00)	97.45 (98.67)	97.45 (99.00)	97.56 (97.67)	97.79 (98.00)	97.12 (98.00)	97.12 (98.00)	97.67 (97.00)	94.13 (96.33)	93.58 (95.67)	98.01 (98.00)
CP	IR-VM	100 (100)	100 (100)	97.55 (94.95)	97.09 (94.50)	96.48 (94.04)	96.79 (94.50)	96.64 (94.50)	96.79 (94.95)	96.79 (94.95)	96.48 (94.50)	96.48 (94.50)	96.94 (94.04)	93.73 (91.28)	93.27 (90.37)	97.25 (95.41)
	IR-WS	100 (100)	100 (100)	97.55 (94.95)	97.09 (94.50)	96.48 (94.04)	96.79 (94.50)	96.64 (94.50)	96.79 (94.95)	96.79 (94.95)	96.48 (94.50)	96.48 (94.50)	96.94 (94.04)	93.73 (91.28)	93.27 (90.37)	97.25 (95.41)

AMV, Alfalfa mosaic virus; MP, movement protein; CP, coat protein.

highest nucleotide identity with isolate HZ (98.09% and 98.04%, respectively). At the individual cistron level, IR-VM and IR-WS shared nucleotide sequence identities of 96.42-99.41% (98.22-99.82% aa sequence identity), 95.19-98.02% (94.05-98.86% aa identity), 93.58-98.78% (95.67-99.33% aa identity), and 93.27-97.55% (90.37-95.41% aa identity) with 13 other isolates in the P1, P2, MP and CP coding regions, respectively (Table 1). AMV-IR-VM and IR-WS shared nucleotide identities of 99.89% (99.67% aa identity), 98.62% (96.33% aa identity) with IR-WS2 (another isolate which only its RNA3 segment obtained from wisteria), in the MP and CP coding regions, respectively. Positions of SNPs in the genome were in P1 (57 SNPs), P2 (101 SNPs), MP (25 SNP), and CP (14 SNPs).

Phylogenetic relationship and reassortment/recombination analysis. The phylogenetic trees were constructed using the NJ method based on each RNA segment (Supplementary Fig. 1) and P1, P2, MP, and CP nucleotide sequences (Fig. 2).

AMV isolates can be divided into two distinct clades, A and B, based on phylogenetic trees of each RNA segment.

Clade A can be further subdivided into two groups. However, some isolates clustered differently (Supplementary Fig. 1). Iranian AMV isolates IR-VM and IR-WS clustered in group A and had a closer genetic relationship to the HZ isolate from China, according to the findings based on RNA1 and RNA3.

Some isolates such as HZ and FERA160224 were assigned to groups I and II respectively, based on the phylogenetic tree of the P1, CP, and MP genomic regions, but were assigned to different groups according to the P2 trees (Fig. 2). Reassortment/recombination may be a possible explanation for this phylogenetic incongruence and differential clustering of these isolates. To identify possible reassortment/recombination events, sequence alignments of RNA1, RNA2, and RNA3 segments of the Iranian isolates along with 13 AMV sequences available from GenBank were concatenated in a single alignment and analyzed using RDP4 implemented algorithms. Any events carrying the warning “this event may not have been caused by recombination” were excluded from the analysis. Six putative reassortment events and one putative recombination event were identified. No significant reassortment/recombination

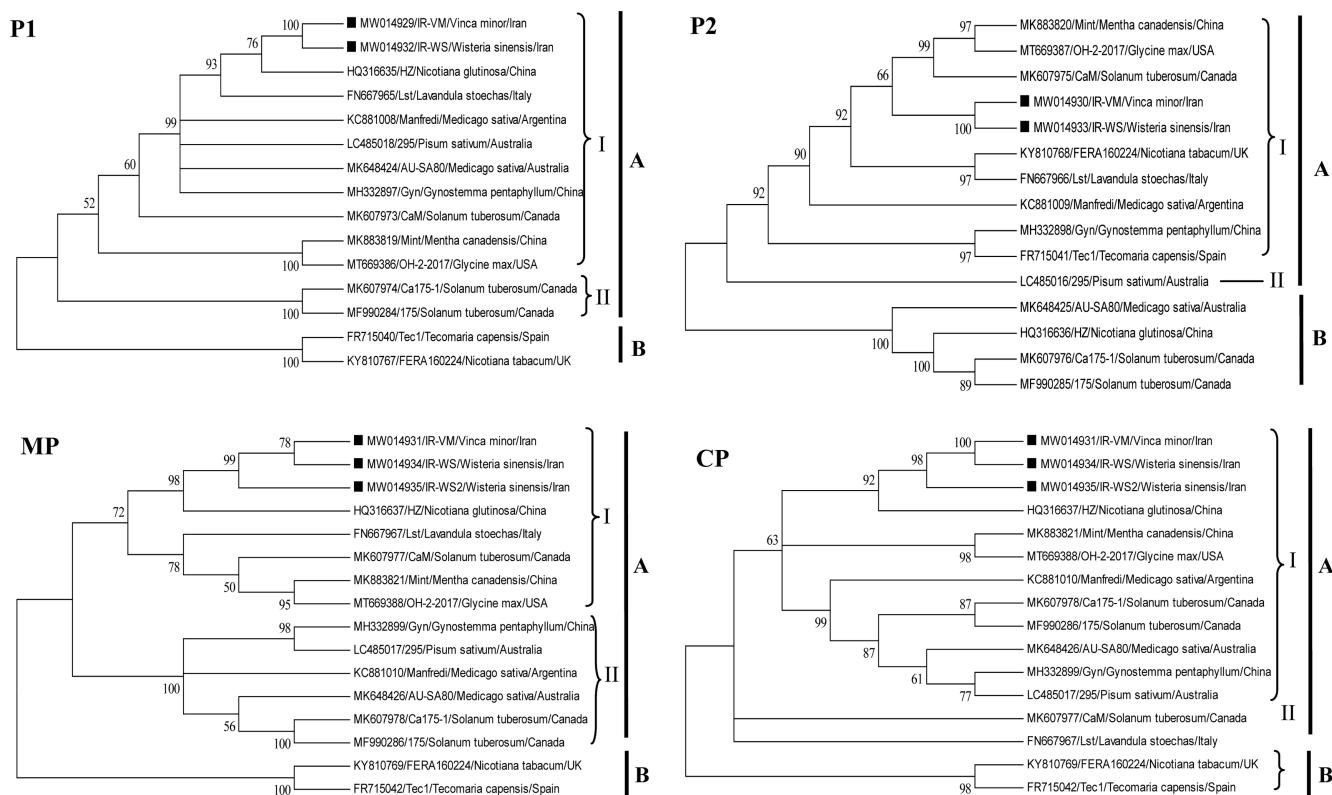


Fig. 2. Phylogenetic trees based on P1, P2, movement protein (MP), and coat protein (CP) nucleotide sequences of Alfalfa mosaic virus (AMV) constructed by the neighbor-joining method using MEGAX, with 1,000 bootstrap replicates. Isolates were indicated in the trees by accession number/isolate name. Bootstrap percentage (BP \geq 50%) are indicated above major branches. Iranian AMV isolate generated from this study was marked.

events were detected in the sequences of Iranian AMV isolates. The information including identification of the reassortment or recombination, position, the major and minor parents, and if other isolates occurred to the same event, the consensus obtained by each of the six different algorithms and the corresponding *P*-values is reported in Supplementary Table 2. In detail, the first putative event HZ resulted in a reassortment of the RNA2 segment (positions 3,696-6,408 nt) which was verified by all algorithms, with Manfredi as major and 175 as minor parental, respectively. Besides, FERA160224 showed a second putative reassortment event in segment RNA2 (pos. 3,542-6,408 nt), confirmed by 6/7 methods, with 295 as major and Lst as minor parental. AU-SA80 and 295 depicted a putative reassortment event in RNA2 (pos. 3,696-6,408 nt), with Gyn and Tec1 sequences as major and minor parental, respectively. Confirmed by 3/7 and 2/7 algorithms, the other two putative reassortment events occurred in the Lst (in the RNA1) and 175 (in the RNA2), respectively. Moreover, Ca175-1 revealed a putative recombinant event in the RNA2 segment (3,759-5,147 nt) with OH-2-2017 as putative a major parent and with CaM as a putative minor parent. This event was verified only in three out of seven algorithms. Phylogenetic analysis further illustrated the evidence of recombination for AMV isolates which is included here (data not shown).

Genetic diversity and selection pressure acting on AMV coding regions. The genetic variation and polymorphism of the AMV isolates were computed for each coding region separately using several genetic diversity parameters implemented in DnaSP6 (Table 2). The order of genetic variation (π) of individual coding regions, from highest to lowest, was as follows: P2, CP, MP, and P1 (Table 2). The *dN/dS* (ω) ratio for each cistron was less than 1, indicating that the AMV genome is under dominant purifying selection. The results showed that the substitutions were not uniformly distributed along the AMV genome. The strongest purify-

ing selection was observed in the P1 protein, supported by the smallest ω value (0.05455), while the weakest purifying selection was in the CP, with the largest ω value (0.18210). Among all the coding regions analyzed here, the highest number of positively selected sites were in P2, suggesting that the P2 may play an important role in AMV evolution (Supplementary Table 3). Pairwise sequence identities were analyzed at both the nucleotide and amino acid levels (data not shown), which confirmed the genetic variation results. Among putative gene products of AMV, CP was the most variable (96.30%), and the P1 (99.08%) were the most conserved proteins. Such results were consistent with genetic polymorphism comparisons.

Phylogenetic analysis and sequence variation in CP gene. No recombination event was found between CP coding region of 107 AMV isolates and the sequences were subjected to phylogenetic analyses. As shown in Fig. 3, all the 107 AMV isolates were divided into two distinct clades: A and B. Isolates of clade A were further divided into three groups (I, II, and III). Group I of clade A is a large and geographically widespread group that included 92 isolates from different parts of the world including Iran (21), Serbia (9), China (7), Croatia (2), Italy (6), Chile (1), USA (16), South Korea (4), Brazil (1), Argentina (1), Australia (9), Canada (7), Czech Republic (1), Bosnia and Herzegovina (1), Egypt (1), England (1) and New Zealand (4). Group II contained two Spanish isolates. Group III consisted of 11 isolates from Mexico (1), France (3), Spain (1), England (3), and New Zealand (3). Clade B contained only two Egyptian isolates. AMV phylogroups did not show a clear division in terms of geographical distribution.

The overall mean value of genetic distance among all AMV isolates was 0.028 ± 0.003 , indicating a high genetic diversity. Group III had the highest genetic distance (0.027 ± 0.004), followed by group I (0.022 ± 0.003), and group II (0.012 ± 0.004). The genetic distance between groups

Table 2. Genetic polymorphism estimated for coding regions of AMV

Genomic region	H	Hd	π	dN	dS	dN/dS ^a
P1 (<i>n</i> = 15)	15	1.000	0.02085	0.00427	0.07827	0.05455
P2 (<i>n</i> = 15)	14	0.990	0.03593	0.01627	0.10735	0.15156
MP (<i>n</i> = 15)	14	0.990	0.03271	0.01110	0.10725	0.10349
CP (<i>n</i> = 15)	13	0.981	0.02908	0.01429	0.07847	0.18210

The π , dS, and dN values were generated according to Jukes and Cantor's method (1969). All statistical analyses were performed using DnaSP6. AMV, Alfalfa mosaic virus; H, number of haplotypes/isolates; Hd, haplotype diversity; π , nucleotide diversity; dS, synonymous nucleotide diversity; dN, nonsynonymous nucleotide diversity; ω = dN/dS, average ratio between nonsynonymous and synonymous substitutions in sequence pairs; MP, movement protein; CP, coat protein.

^aMean (dN/dS) value = 1, <1, and >1 indicate neutral evolution, negative (purifying) selection and positive (diversifying) selection, respectively, for each gene-specific sequence data set.

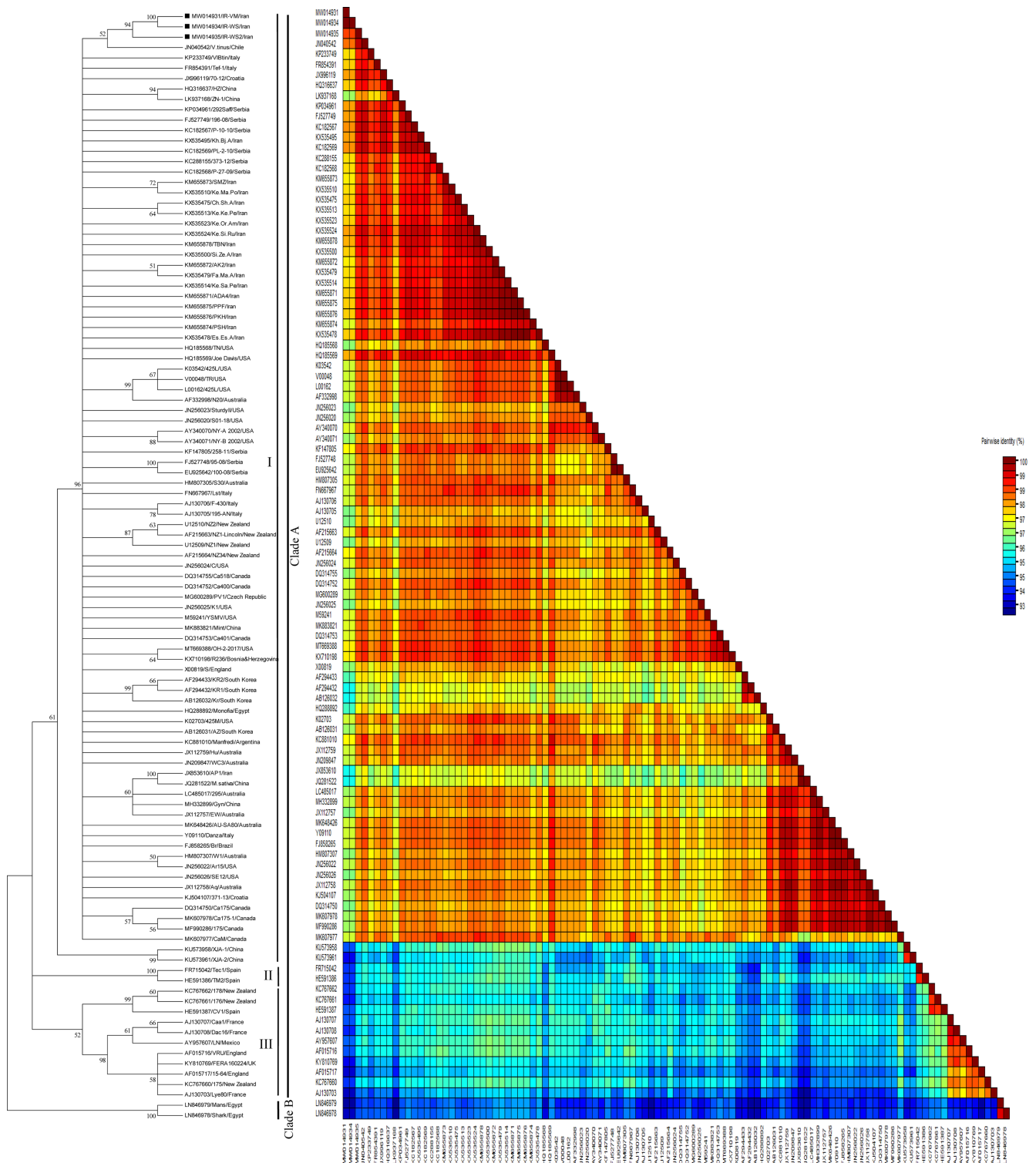


Fig. 3. Neighbor-joining phylogenetic tree constructed from the coat protein gene nucleotide sequences of 107 Alfalfa mosaic virus isolates (three from this study and 104 from the GenBank database), and graphical representation of pairwise nucleotide identity (with percentage identity scale). The phylogenetic tree was generated in MEGAX and bootstrapped with 1,000 replicates. Isolates are indicated in the trees by accession number/isolate name/geographical origin of collection. Bootstrap values $\geq 50\%$ are shown at the branch internodes. Two-dimensional nucleotide diversity plot constructed based on SDT MUSCLE alignment.

Table 3. Summary of population genetics parameters and neutrality tests calculated for the CP gene sequences of AMV population

Variant group	H	Hd	S	η	K	π	dN	dS	ω	Tajima's D	Fu and Li's D*	Fu and Li's F*
Total (<i>n</i> = 107)	96	0.997 (\pm 0.002)	202	226	18.896	0.02889 (\pm 0.00195)	0.01444	0.08041	0.17957	-1.87522*	-3.11329*	-3.08050**
Group I (<i>n</i> = 92)	81	0.996 (\pm 0.003)	158	173	13.853	0.02118 (\pm 0.00124)	0.00965	0.06081	0.15869	-1.99578*	-2.73987*	-2.90847*
Group II (<i>n</i> = 2)	2	1.000 (\pm 0.500)	8	8	8	0.01207 (\pm 0.00603)	0.00600	0.03170	0.18927	NA	NA	NA
Group III (<i>n</i> = 11)	11	1.000 (\pm 0.039)	56	59	17.745	0.02677 (\pm 0.00309)	0.01451	0.06978	0.20793	-0.56472 ^{ns}	-0.72851 ^{ns}	-0.77930 ^{ns}
Geographic origin												
Iranian isolates (<i>n</i> = 21)	17	0.967 (\pm 0.030)	48	50	7.962	0.01217 (\pm 0.00270)	0.00756	0.02771	0.27282	-1.70130 ^{ns}	-1.71187 ^{ns}	-1.99536 ^{ns}
East Asia iso- lates (<i>n</i> = 11)	11	1.000 (\pm 0.039)	72	75	23.127	0.03536 (\pm 0.00478)	0.01624	0.10518	0.15440	-0.46166 ^{ns}	-0.04882 ^{ns}	-0.17599 ^{ns}
European iso- lates (<i>n</i> = 36)	35	0.998 (\pm 0.007)	121	132	22.165	0.03389 (\pm 0.00248)	0.01672	0.09570	0.17471	-1.13604 ^{ns}	-2.16163 ^{ns}	-2.14379 ^{ns}
American iso- lates (<i>n</i> = 27)	24	0.989 (\pm 0.015)	79	81	13.900	0.02125 (\pm 0.00218)	0.00963	0.06115	0.15748	-1.30858 ^{ns}	-2.48030 ^{ns}	-2.47559 ^{ns}
Oceania isolates (<i>n</i> = 9)	9	1.000 (\pm 0.052)	27	27	7.722	0.01165 (\pm 0.00283)	0.00478	0.03448	0.13863	-1.11186 ^{ns}	-0.97346 ^{ns}	-1.12933 ^{ns}
African isolates (<i>n</i> = 3)	3	1.000 (\pm 0.272)	47	47	31.333	0.04726 (\pm 0.01998)	0.03473	0.09935	0.34957	NA	NA	NA

All statistics were performed using DnaSP version 6.10.04.

CP, coat protein; AMV, Alfalfa mosaic virus; H, number of haplotypes; Hd, haplotype diversity; S, number of polymorphic (segregating) sites; η (eta), total number of mutations; K, average number of nucleotide differences between sequences; π , nucleotide diversity (average nucleotide substitutions per site between sequence pairs), with Jukes & Cantor correction; dN, average number of nonsynonymous substitutions per nonsynonymous site; dS, average number of synonymous substitutions per synonymous site, with the Jukes and Cantor correction; dN/dS, average ratio between nonsynonymous and synonymous substitutions in sequence pairs; NA, not available due to limited sequences for analysis. * $P < 0.05$; ** $P < 0.02$; ns, not significant ($P > 0.10$).

was higher than within-subgroups (Supplementary Table 4) which supports evidence for the phylogenetic grouping.

AMV-IR-VM, IR-WS, and IR-WS2 shared 92.66-99.54% nt identity with those of 104 other AMV isolates available in the GenBank, the highest with isolate V.tinus (JN040542) from Chile (with 98.17%, 98.17%, and 99.54% nt identity respectively), and the lowest with isolate Mans (LN846979) from Egypt (with 92.66%, 92.66%, and 93.73% nt identity respectively). The obtained isolates also shared 88.99-99.08% aa identity with other AMV isolates retrieved from GenBank, indicating most of the CP nucleotide substitutions were nonsynonymous with no change in the physicochemical properties of the protein. Haplotype and nucleotide diversity for all AMV isolates were 0.997 and 0.02889, respectively, indicating high haplotype diversity and relatively low nucleotide diversity in AMV populations and among lineage subpopulations (Table 3). The haplotype diversity for groups I, II, and III

was 0.996, 1.000, and 1.000, whereas nucleotide diversity for these three groups was 0.02118, 0.01207, and 0.02677, respectively. The higher overall average number of nucleotide differences ($k = 17.745$) was calculated for phylogroup III. However, the greater number of segregation sites, ($S = 158$), and mutations within the segregating sites, ($\eta = 173$), were found in the phylogroup I (Table 3).

At the geographical population level, the highest and lowest values of π and k were estimated for the African ($\pi = 0.04726$, $k = 31.333$) and Oceania ($\pi = 0.01165$, $k = 7.722$) populations, respectively. However, the highest values of S (121), and η (132), were found in European isolates, while the values of S (27), and η (27) in Oceania isolates were the lowest. The global selection pressure (dN/dS) in CP of all AMV isolates was 0.17957. Furthermore, the dN values for each population were less than the dS values (dN/dS ratio < 1), suggested purifying selection was restricting variability in the population. The highest and lowest dN/dS

Table 4. Codon positions of CP coding region of 107 AMV isolates significantly affected by positive selection by different codon-based maximum-likelihood algorithms

Model	Positively selected sites
SLAC	176, 212
FEL	133, 150, 176, 212, 221
IFEL	87, 94, 133, 176, 212
FUBAR	133, 176, 212, 221
MEME	9, 11, 15, 104, 145, 150, 176, 185, 212, 213

In bold are the sites identified as being under positive selection by more than three methods.

Significance levels set at $P = 0.1$ for SLAC, FEL, IFEL, and MEME programs, posterior probability of 0.9 for FUBAR.

CP, coat protein; AMV, Alfalfa mosaic virus; SLAC, single-likelihood ancestor counting; FEL, fixed effects likelihood; IFEL, internal fixed effects likelihood. All statistics were performed using DnaSP version 6.10.04; MEME, mixed effects model of evolution.

ratio was calculated for African ($\omega = 0.34957$) and Oceania ($\omega = 0.13863$) populations. Quantifying selection pressure at individual codon positions showed that most of the codons were under negative selection or neutral evolution, while the codons at positions 133, 176, and 212 were found under positive selection by at least three methods imple-

mented in HYPHY. Notably, among these codon positions, only the amino acid position 212 (in the C-terminal region of CP) was confirmed by all the five methods (Table 4). The CP N-terminal domain has been shown to be involved in cell-to-cell and systemic movement in members of the different genera within the family *Bromoviridae* including AMV (Sánchez-Navarro and Bol, 2001). It has been demonstrated that N-terminal K residues at positions 5 to 13 within the K5KAGGKAGK13 sequence are essential and sufficient for transportation of AMV CP to the nucleolus (Herranz et al., 2012). Some of the changes, for example, replacement of underline amino acid from A (nonpolar) to S (polar) with different physicochemical properties were observed in the CP of Iranian isolates. In addition, the Alanine to Serine mutation can modify protein formation due to hydrophobic (Alanine) to hydrophilic (Serine) molecule changes. The specified effect of the mentioned mutations in this stretch should be further studied.

Genetic differentiation and population structure of AMV based on CP sequences. The genetic distinction of AMV populations was defined in two categories: phylogenetic populations and geographical populations. Pairwise

Table 5. Genetic differentiation measurement between subpopulations from pairwise comparison of AMV CP sequences

Comparisons	K_S^*	K_{ST}^*	P -value	Z^*	P -value	S_{nn}	P -value	F_{ST}
Phylogroup								
Group I ($n = 92$) vs. II ($n = 2$)	2.59292	0.01423	0.0020**	7.34522	0.0000***	1.00000	0.0020**	0.65645
Group I ($n = 92$) vs. III ($n = 11$)	2.61212	0.05764	0.0000***	7.38353	0.0000***	1.00000	0.0000***	0.52664
Group II ($n = 2$) vs. III ($n = 11$)	2.83366	0.05022	0.0110*	3.13784	0.0090**	1.00000	0.0170*	0.55472
Geographic origin								
Iran ($n = 21$) vs. East Asia ($n = 11$)	2.29027	0.08618	0.0000***	4.95190	0.0000***	0.85938	0.0000***	0.17205
Iran ($n = 21$) vs. Europe ($n = 36$)	2.62338	0.05059	0.0000***	6.13135	0.0000***	0.83333	0.0000***	0.16815
Iran ($n = 21$) vs. America ($n = 27$)	2.30127	0.06243	0.0000***	5.74973	0.0000***	0.83333	0.0000***	0.15674
Iran ($n = 21$) vs. Oceania ($n = 9$)	1.94013	0.12933	0.0000***	4.70367	0.0000***	0.90000	0.0000***	0.39970
Iran ($n = 21$) vs. Africa ($n = 3$)	1.98377	0.12971	0.0010**	4.43907	0.0040**	0.95833	0.0070**	0.39342
East Asia ($n = 11$) vs. Europe ($n = 36$)	3.02492	0.01609	0.0040**	5.88543	0.0030**	0.89362	0.0000***	0.10069
East Asia ($n = 11$) vs. America ($n = 27$)	2.71435	0.02162	0.0080**	5.43570	0.0090**	0.84211	0.0010**	0.06119
East Asia ($n = 11$) vs. Oceania ($n = 9$)	2.59222	0.04586	0.0030**	4.08754	0.0040**	0.78333	0.0080**	0.12690
East Asia ($n = 11$) vs. Africa ($n = 3$)	3.07074	0.05652	0.0240*	3.25615	0.0210*	0.92857	0.0280*	0.27365
Europe ($n = 36$) vs. America ($n = 24$)	2.83389	0.02052	0.0020**	6.46465	0.0000***	0.74471	0.0010**	0.09851
Europe ($n = 36$) vs. Oceania ($n = 9$)	2.83871	0.03738	0.0000***	5.67525	0.0000***	0.83704	0.0090**	0.26428
Europe ($n = 36$) vs. Africa ($n = 3$)	3.01769	0.02268	0.0010**	5.49569	0.0020**	0.97436	0.0070**	0.23637
America ($n = 24$) vs. Oceania ($n = 9$)	2.45636	0.01586	0.0460*	5.38797	0.0450*	0.76852	0.0170*	0.12691
America ($n = 24$) vs. Africa ($n = 3$)	2.60906	0.05246	0.0000***	4.91621	0.0000***	0.96667	0.0010**	0.33001
Oceania ($n = 9$) vs. Africa ($n = 3$)	2.12975	0.19328	0.0050**	2.82153	0.0070**	0.91667	0.0350*	0.41676

Probability (P -value) obtained by the permutation test (PM test) with 1,000 replicates. * $0.01 < P < 0.05$; ** $0.001 < P < 0.01$; *** $P < 0.001$; ns, not significant. The PM test was performed using DnaSP version 6.10.04. $F_{ST} > 0.33$ indicates infrequent gene flow; $F_{ST} < 0.33$ suggests frequent gene flow. Any comparison with Group IV was not available.

AMV, Alfalfa mosaic virus; CP, coat protein.

F_{ST} values showed a strong genetic differentiation ($F_{ST} > 0.52664$) between phylogroups of AMV isolates. Further support for these results was found in significantly high values of Ks^* , Kst^* , Z^* , and Snn statistics (Table 5) displayed that the isolates between phylogroups had a very high genetic differentiation and infrequent gene flow ($F_{ST} > 0.33$), which is consistent with the phylogenetic analysis in Fig. 3. Multiple lineages composed AMV populations from different geographic origins, among which significant high Ks^* , Kst^* , Z^* , and Snn values and great F_{ST} value (over 0.25) were obtained between the AMV populations from Iran, East Asia, and Europe with Oceania and Africa as shown in Table 5. This suggested complete genetic differentiation and infrequent gene flow. A very high degree of differentiation was observed between America vs. Africa, and Oceania vs. Africa confirmed by high F_{ST} value (> 0.33) and significant Ks^* , Kst^* , Z^* , and Snn values. The level of genetic differentiation between Iran vs. Europe, East Asia, and Africa was great, as revealed by $F_{ST} = 0.15674-0.17205$. Genetic differentiation between East Asia vs. America, Europe, and Oceania populations and America vs. Oceania populations confirmed by moderate levels of F_{ST} (0.06119-0.09851). In most populations, there were moderate levels of gene flow despite large separating geographic distances.

Neutrality tests. The values of Tajima's D, Fu and Li's D^* & F^* statistics were negative and significant for group I and the overall population suggested that the estimated polymorphism was less than what expected. However, the three indices were negative and statistically insignificant ($P > 0.10$) for group III. A negative value of Tajima's D signifies an excess of low-frequency polymorphisms caused by background selection, genetic hitchhiking, or population expansions (Tajima, 1989). No significant departure from neutrality was found for any geographical groups of AMV populations when the CP sequences were tested for Tajima's D, Fu and Li's D^* and F^* values; nevertheless, the values were negative for all populations.

Discussion

In this study, we reported the first complete genome sequences of AMV isolated from *V. minor* and *W. sinensis*, assembled by RNA-seq reads and RT-PCR. To the best of our knowledge, this is the first confirmed report of AMV infecting dwarf periwinkle and wisteria in the world. Pairwise identity analysis showed relatively significant variability in genome sequences among AMV isolates with genomic nucleotide identities of 94.78-99.97 (average, 97.42%), 92.54-99.55 (average, 95.86%), and 93.89-

99.95 (average, 97.11%) in RNA1, RNA2, and RNA3, respectively. Although genetic polymorphism analysis for each coding region showed low nucleotide diversity, it also displayed high genetic diversity following the high haplotype diversity, which is a signature of a population expansion with little selection on any particular gene. Of all the coding regions analyzed, the P2 gene indicated higher nucleotide diversity ($\pi = 0.03593$). Apart from codon mutations, insertions/deletions also contributed to genetic variability in the P2 protein. Different selective constraints imposed on different AMV proteins might be associated with the diverse functions of these proteins in the infection cycle of the virus and/or their interactions with the host and aphid vectors (García-Arenal et al., 2001). The evolutionary constraints enforced on proteins P1, P2, and MP were larger than those exerted on protein CP, indicating a greater tolerance for amino acid substitutions among the CP gene. Although most codons were under negative selection or neutral evolution, a few codons were identified under positive selection in P1 and P2 cistrons by at least two separate algorithms within the HYPHY software package (Supplementary Table 3). Positive selection on some codons in AMV coding regions confirmed a molecular adaptation by fixing beneficial genotypes and enhancing population diversification. This finding also presents additional support for the wide spectrum of AMV infection. Our analyses distinguished six putative reassortment events and one putative recombination event among some AMV isolates. These events point to the importance of this phenomenon in the evolution of global AMV populations, which is in line with that of other investigations (Bergua et al., 2014; Bonnet et al., 2005). This shows that genetic exchange by recombination or reassortment transpires naturally in populations of plant viruses with segmented genomes. It is expected that the origins and evolution of AMV be investigated to highlight the organization and structure of its genetic diversity and the role of the evolutionary forces that have been shaping this diversity.

In this research, the CP gene of 107 AMV isolates was used to investigate the genetic diversity and molecular mechanisms underlying the evolution of this virus. The 107 AMV isolates fell into two main clades which have significant subpopulation differentiation with an F_{ST} value of 0.71959. Higher haplotype diversity and low nucleotide diversity indicated that the AMV population studied recently diverged from each other.

The observed significant genetic differentiation and low gene flow in some populations suggest that geographical isolation would be a barrier to gene flow. However, moderate levels of genetic differentiation and gene flow were ob-

served between most of geographical populations, indicating the few influences of temporal and spatial variations on population differentiation in these populations. Such gene flow can potentially reduce genetic differentiation. Taken together, long-distance relocation, most likely by human transport of contaminated seeds or propagative plant material, has caused genetically related virus isolates to be detected in far-off geographical districts. This outcome is in concurrence with Bergua et al. (2014). The presence of diversifying selection at positions 133, 176, and 212 in the AMV CP coding region, reflects the adaptive evolution. The AMV multifunctional CP has significant roles in plus-strand RNA accumulation, pathogenicity, virion assembly (Houwing and Jaspars, 1993), vector specificity, regulation of replication and translation of viral RNAs, cell-to-cell movement (Sánchez-Navarro and Bol, 2001; van der Vossen et al., 1994) and systemic spread of the virus (Bol 2008; Neeleman et al., 2004; Tenllado and Bol, 2000). An SNP A→T (26.70% variant frequency) occurred in codon position 133 (at nucleotide 399) and represented the polymorphism that led to a change in the encoded amino acid from glutamine (Q, codon CAA) to histidine (H, codon CAT). It is not known if the identified amino acid change in the central region of CP would affect the virus's life cycle. Two domains located at the N- and C-terminal ends of AMV CP are involved in its nuclear import and export, respectively. Interestingly, two nonpolar amino acids, leucine (L, codon CTC), and phenylalanine (F, codon TTC), at codon position 212 were observed to be under positive selection. Sequence alignment showed that most of the sequences encoded the amino acid leucine at codon position 212 while some others encoded phenylalanine. It has also been previously shown that mutations at the N and C termini affect the movement of viral materials through the vascular system (Tenllado and Bol, 2000). The significance of these positive selections on the codons at positions 133, 176, and 212 in the host-virus–vector interactions need to be investigated further.

Significantly negative Tajima's D, Fu and Li's D* & F* values for group I and overall population suggesting the principle of operation of purifying selection and population size expansion. For all geographical groups, all three neutrality tests were negative, which might be the result of population expansion/recovery from a population bottleneck. However, since the *P*-value was not significant, the result was not conclusive, and all populations seem to be at drift-mutation equilibrium.

Iran has the optimal conditions for transmission such as aphids. The expansive spread of AMV in Iran can cause problems in different regions of the country via its aphid

vectors. Therefore, it is important to use cultivars resistant to virus. Vinca and wisteria are perennial plants and cultivated in Iran as medical and ornamental plants. They could serve as a potential reservoir for AMV to infect other ornamentals and cultivated crops (Bergua et al., 2014). Moreover, AMV can easily be transmitted through vegetative propagation of wisteria (by cuttings and grafting) and dwarf periwinkle (by stem cuttings) and dispersed. This proposes that strict quarantine regulation is necessary to prevent the movement of novel variants or allelic combinations of AMV between the various regions in the country when trading plant materials that are hosts for this pathogen.

To sum up, the current study represents an integrated analysis of the population genetic structure of AMV based on CP gene sequences. The phylogenetic analysis did not show a clear geography-specific clustering of AMV isolates. Gene flow, recombination/reassortment, and negative selection were found to be the important evolutionary factors that can affect the genetic structure of AMV populations. Founder effects by exchanging infected plant material between different geographical regions have also been suggested to play a role in shaping the genetic structure of AMV. Our findings provide basic information for the development of an integrated disease management strategy against AMV in Iran.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This work was supported by Sari Agricultural Sciences and Natural Resources University (No. 01-1400-03).

Electronic Supplementary Material

Supplementary materials are available at The Plant Pathology Journal website (<http://www.ppjonline.org/>).

References

- Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M. and Boonham, N. 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Mol. Plant Pathol.* 10:537-545.
- Al-Shahwan, I. M. 2002. Alfalfa mosaic virus (AMV) on alfalfa (*Medicago sativa* L.) in Saudi Arabia. *Assiut J. Agric. Sci.*

- 33:21-30.
- Bergua, M., Luis-Arteaga, M. and Escriu, F. 2014. Genetic diversity, reassortment, and recombination in Alfalfa mosaic virus population in Spain. *Phytopathology* 104:1241-1250.
- Bol, J. F. 1999. Alfalfa mosaic virus and ilarviruses: involvement of coat protein in multiple steps of the replication cycle. *J. Gen. Virol.* 80:1089-1102.
- Bol, J. F. 2008. *Alfalfa Mosaic Virus*. In: *Encyclopedia of virology*, 3rd ed., eds. by B. W. J. Mahy and M. H. V. van Regenmortel, pp. 81-87. Academic Press, Oxford, UK.
- Bonnet, J., Fraile, A., Sacristán, S., Malpica, J. M. and García-Arenal, F. 2005. Role of recombination in the evolution of natural populations of Cucumber mosaic virus, a tripartite RNA plant virus. *Virology* 332:359-368.
- Bujarski, J., Figlerowicz, M., Gallitelli, D., Roossinck, M. J. and Scott, S. W. 2012. Family *Bromoviridae*. In: *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, eds. by A. M. Q. King, M. J. Adams, E. B. Carstens and E. J. Lefkowitz, pp. 965-976. Academic Press, Oxford, UK.
- Che, X., Jiang, X., Liu, X., Luan, X., Liu, Q., Cheng, X. and Wu, X. 2020. First report of Alfalfa mosaic virus on soybean in Heilongjiang, China. *Plant Dis.* 104:3085.
- Cuevas, J. M., Delaunay, A., Rupar, M., Jacquot, E. and Elena, S. F. 2012. Molecular evolution and phylogeography of potato virus Y based on the CP gene. *J. Gen. Virol.* 93:2496-2501.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792-1797.
- Edwardson, J. R. and Christie, R. G. 1997. *Alfamovirus* Genus. *Alfalfa mosaic virus* species. In: *Viruses infecting peppers and other solanaceous crops*, eds. by J. R. Edwardson and R. G. Christie, pp. 63-94. University of Florida Press, Gainesville, FL, USA.
- Esfandiari, N., Kohi Habibi, M., Mosahebi, G. H. and Mozafari, J. 2005. Detection of Alfalfa mosaic virus (AMV) in pea field in Iran. *Commun. Agric. Appl. Biol. Sci.* 70:407-410.
- Farahanikia, B., Akbarzadeh, T., Jahangirzadeh, A., Yassa, N., Shams Ardekani, M. R., Mirnezami, T., Hadjiakhoondi, A. and Khanavi, M. 2011. Phytochemical investigation of *Vinca minor* cultivated in Iran. *Iran. J. Pharm. Res.* 10:777-785.
- Fletcher, J. D. 2001. New hosts of Alfalfa mosaic virus, Cucumber mosaic virus, Potato virus Y, Soybean dwarf virus, and Tomato spotted wilt virus in New Zealand. *N. Z. J. Crop Hort. Sci.* 29:213-217.
- Fu, Y. X. and Li, W. H. 1993. Statistical tests of neutrality of mutations. *Genetics* 133:693-709.
- Gao, F., Lin, W., Shen, J. and Liao, F. 2016. Genetic diversity and molecular evolution of arabis mosaic virus based on the CP gene sequence. *Arch. Virol.* 161:1047-1051.
- García-Arenal, F., Fraile, A. and Malpica, J. M. 2001. Variability and genetic structure of plant virus populations. *Annu. Rev. Phytopathol.* 39:157-186.
- Golnaraghi, A. R., Shahraeen, N., Pourrahim, R., Farzadfar, S. and Ghasemi, A. 2004. Occurrence and relative incidence of viruses infecting soybeans in Iran. *Plant Dis.* 88:1069-1074.
- Hamzeh, N., Koochi Habibi, M., Mosahebi, G., Dizadji, A. and Ghazanfari, K. 2010. Occurrence of Tomato spotted wilt virus, Cucumber mosaic virus and Alfalfa mosaic virus in Narcissus an ornamental plant in Iran. In: *19th Iranian Plant Protection Congress*. Tehran, Iran.
- He, B., Fajolu, O. L., Wen, R.-H. and Hajimorad, M. R. 2010. Seed transmissibility of Alfalfa mosaic virus in soybean. *Plant Health Prog.* 11:41.
- Herranz, M. C., Pallas, V. and Aparicio, F. 2012. Multifunctional roles for the N-terminal basic motif of Alfalfa mosaic virus coat protein: nucleolar/cytoplasmic shuttling, modulation of RNA-binding activity, and virion formation. *Mol. Plant-Microbe Interact.* 25:1093-1103.
- Hiruki, C. and Hampton, R. O. 1990. Diseases caused by viruses and viruses infectious to alfalfa. In: *Compendium of Alfalfa diseases*, 2nd ed., eds. by D. L. Stuteville and D. C. Erwin, pp. 51-58. American Phytopathological Society, St. Paul, MN, USA.
- Houwing, C. J. and Jaspars, E. M. 1993. Coat protein stimulates replication complexes of Alfalfa mosaic virus to produce virion RNAs *in vitro*. *Biochimie* 75:617-621.
- Hudson, R. R. 2000. A new statistic for detecting genetic differentiation. *Genetics* 155:2011-2014.
- Jasper, E. M. J. and Bos, L. 1980. Alfalfa mosaic virus. Association of Applied Biologists Description of Plant Viruses N8 229. URL <http://www.dpvweb.net/dpv/showdpv.php?dpvno=229> [3 November 2021].
- Jiang, Y., Chen, X., Lin, H., Wang, F. and Chen, F. 2011. Floral scent in wisteria: chemical composition, emission pattern, and regulation. *J. Am. Soc. Hort. Sci.* 136:307-314.
- Jukes, T. H. and Cantor, C. R. 1969. Evolution of protein molecules. In: Munro HN (ed) *Mammalian protein metabolism*, ed. by H. N. Munro, pp. 21-132. Academic Press, New York, USA.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35:1547-1549.
- Maina, S., Zheng, L., Kinoti, W. M., Aftab, M., Nancarrow, N., Trębicki, P., King, S., Constable, F. and Rodoni, B. 2019. Metagenomic analysis reveals a nearly complete genome sequence of Alfalfa mosaic virus from a field pea in Australia. *Microbiol. Resour. Announc.* 8:e00766-19.
- Mangeli, F., Massumi, H., Alipour, F., Maddahian, M., Heydarnejad, J., Hosseinipour, A., Amid-Motlagh, M. H., Azizizadeh, M. and Varsani, A. 2019. Molecular and partial biological characterization of the coat protein sequences of Iranian Alfalfa mosaic virus isolates. *J. Plant Pathol.* 101:735-742.
- Martin, D. P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B. 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* 1:vev003.
- Massumi, H., Maddahian, M., Heydarnejad, J., Hosseini Pour, A. and Farahmand, A. 2012. Incidence of viruses infecting al-

- falfa in the southeast and central regions of Iran. *J. Agric. Sci. Technol.* 14:1141-1148.
- Moradi, Z. and Mehrvar, M. 2019. Genetic variability and molecular evolution of Bean common mosaic virus populations in Iran: comparison with the populations in the world. *Eur. J. Plant Pathol.* 154:673-690.
- Neeleman, L., Linthorst, H. and Bol, J. F. 2004. Efficient translation of alfamovirus RNAs requires the binding of coat protein dimers to the 3' termini of the viral RNAs. *J. Gen. Virol.* 85:231-240.
- Pond, S. L. K. and Frost, S. D. W. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21:2531-2533.
- Pourrahim, R. and Farzadfar, S. 2015. Biological and molecular characterization of *Alfalfa mosaic virus* infecting trumpet creeper (*Campsis radicans*) in Iran. *J. Phytopathol.* 164:276-280.
- Prabha, K., Baranwal, V. K. and Jain, R. K. 2013. Applications of next generation high throughput sequencing technologies in characterization, discovery and molecular interaction of plant viruses. *Indian J. Virol.* 24:157-165.
- Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E. and Sanchez-Gracia, A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34:3299-3302.
- Sánchez-Navarro, J. A. and Bol, J. F. 2001. Role of the Alfalfa mosaic virus movement protein and coat protein in virus transport. *Mol. Plant-Microbe Interact.* 14:1051-1062.
- Sawalha, H. and Mansour, A. 1996. Incidence of Alfalfa mosaic virus in alfalfa fields in Jordan. *Derasat* 23:81-83.
- Smit, C. H. and Jaspars, E. M. 1982. Evidence that RNA 4 of Alfalfa mosaic virus does not replicate autonomously. *Virology* 117: 271-274.
- Song, S., Liu, H., Zhang, J., Pan, C. and Li, Z. 2019. Identification and characterization of complete genome sequence of Alfalfa mosaic virus infecting *Gynostemma pentaphyllum*. *Eur. J. Plant Pathol.* 154:491-497.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- Tenllado, F. and Bol, J. F. 2000. Genetic dissection of the multiple functions of Alfalfa mosaic virus coat protein in viral RNA replication, encapsidation, and movement. *Virology* 268:29-40.
- Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.
- United States Department of Agriculture. 2020. *Vinca minor* L.: common periwinkle. URL <https://plants.usda.gov/home/plantProfile?symbol=VIMI2> [3 November 2021].
- van der Vossen, E. A. G., Neeleman, L. and Bol, J. F. 1994. Early and late functions of Alfalfa mosaic virus coat protein can be mutated separately. *Virology* 202:891-903.
- van Dun, C. M., Bol, J. F. and Van Vloten-Doting, L. 1987. Expression of Alfalfa mosaic virus and tobacco rattle virus coat protein genes in transgenic tobacco plants. *Virology* 159:299-305.
- Vas, A. and Gulyás, B. 2005. Eburnamine derivatives and the brain. *Med. Res. Rev.* 25:737-757.
- Weimer, J. L. 1931. Alfalfa mosaic virus. *Phytopathology* 21:122-123.
- Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* 15:323-354.
- Wu, Q., Ding, S.-W., Zhang, Y. and Zhu, S. 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. *Annu. Rev. Phytopathol.* 53:425-444.
- Xu, H. and Nie, J. 2006. Identification, characterization, and molecular detection of *Alfalfa mosaic virus* in Potato. *Phytopathology* 96:1237-1242.
- Zenaddini, A., Jafarpour, B. and Falahati Rastegar, M. 2004. Identification and study on properties and distribution of Alfalfa mosaic virus in Khorasan province. In: *19th Iranian Plant Protection Congress*. Tabriz, Iran.