Analyses of complete nucleotide sequence of Iranian isolate of maize dwarf mosaic virus (MDMV) and notes on the origin and evolution of MDMV*

F. S. MOSTAFAVI NEISHABURI[1], M. MASUMI[2], S. NASROLLANEJAD[3], N. RAHPEYMA-SARVESTANI[4] and K. IZADPANAH[5]**

(Received: 21.1.2014; Accepted: 18.1.2015)

Abstract

Complete nucleotide sequence of an Iranian isolate of Maize dwarf mosaic virus (MDMV-Ir) was determined. The viral genome comprises 9499 nucleotides, excluding a 3′-terminal poly A tail. It consists of a 139 nucleotide long 5′noncoding region, a 234 nucleotide long 3′noncoding region and an open reading frame (ORF) which encodes a single polyprotein of 3042 amino acid residues. MDMV-Ir shares 86.4% to 91.9% nucleotide identity and 93.9% to 97.3% amino acid sequence similarity with other isolates of the virus. The nucleotide sequence identity of MDMV-Ir with full sequences of other cereal potyviruses available in GenBank ranged from 54 to 70.5%. The putative amino acid sequences of MDMV-Ir proteins showed 29.8% to 60% similarity to corresponding proteins of other cereal viruses in the GenBank. Sequence comparisons and phylogenetic analyses of the complete genome and derived polyproteins revealed that MDMV-Ir has the highest similarity to the Bulgarian isolate of MDMV, and among other cereal potyviruses it has the closest relationship to Sorghum mosaic virus. These analyses also indicated that MDMV populations are a monophyletic group originated in the Mediterranean basin.

Keywords: *Potyviridae*, Maize dwarf mosaic virus, Mediterranean Basin, Virus origin, Iran

* A part of M.Sc. thesis submitted by the first author in Faculty of Agriculture Sciences, Shahrekord University, Shahrekord, Iran.

** Corresponding Author, Email: masoumi@shirazu.ac.ir

1. Graduate student of Plant Pathology, Gorgan University of Agriculture and Natural Resources. Gorgan, Iran
2. Research Assistant Professor, Plant Virology Research Center, Shiraz University and Agricultural Research and Education Organization, Agricultural and Natural Resources Research Center of Fars. Shiraz, Iran
3. Associate Professor, Gorgan University of Agriculture and Natural Resources, Dept. of Plant Protection. Gorgan, Iran
4. Graduate student of Plant Breeding, Shiraz University, Shiraz, Iran
5. Professor of Plant Pathology, Plant Virology Research Center, Shiraz University, Shiraz, Iran

Introduction

The genus *Potyvirus* (family *Potyviridae*) comprises the most economically important plant viruses (Hollings and Brunt 1981). Potyviruses infecting poaceous plants are often referred to as *Sugarcane mosaic virus* (SCMV) subgroup. The latter, in addition to SCMV, include *Maize dwarf mosaic virus* (MDMV), *Sorghum mosaic virus* (SrMV), *Pennisetum mosaic virus* (PenMV), *Zea mosaic virus* (ZeMV), *Johnson grass mosaic virus* (JGMV), Iranian Johnson grass mosaic virus (IJMV), newly reported Bermuda grass southern mosaic virus (BgSMV) and reed canary grass mosaic virus (RCGMV) (Fan *et al.* 2003; Fan and Li 2004; Heidary *et al.* 2008; Masumi *et al.* 2001; 2004; Masumi and Izadpanah 1998; Seifers *et al.* 2000; Salomon and Seifers 2004; Zare *et al.* 2005; 2006). Most of these viruses can infect maize. They all have a single stranded RNA genome of approximately 9500 nucleotides, encoding a single large polyprotein that is proteolytically processed by the virus encoded proteases yielding 10 functional proteins (Shukla and Ward 1988). To date, the complete sequences of more than 20 isolates of SCMV-subgroup have been reported and deposited in database. Only five isolates of MDMV from Bulgaria (MDMV-Bg), Spain (MDMV-Sp), Italy (MDMV-It) and USA (MDMV-OH1, MDMV-OH2) have been completely sequenced (Achon *et al.* 2007; Kong and Steinbiss 1998; Stewart *et al.* 2012).

Both MDMV and IJMV cause maize mosaic in Iran. While IJMV is distributed, with Johnson grass (*Sorghum halepense* (L.) Pers.), all over Iran, MDMV has been reported mainly from northern and to a lower extent from central regions of the country (Masumi *et al.* 2004). Serological and biological properties of MDMV-Ir in comparison with other cereal potyviruses in Iran have been investigated by Zare *et al.* 2004. Similar studies have been made on nucleotide sequence of 5′-region of CP gene of the virus. Here, we analyze the complete genome sequence of MDMV-Ir to show its evolutionary and phylogenetic relationship with other isolates of the virus in the Mediterranean basin. We also use available information to speculate on possible origin of MDMV.

Materials and Methods

Virus isolation, purification and genome amplification

Maize plants with mosaic and dwarfing symptoms were collected from fields of Golestan province, Iran, and their infection by MDMV was verified by indirect ELISA (Converse and Martin 1990) using an antiserum against this virus (Zare *et al.* 2004). The virus was isolated and mechanically transmitted to *Sorghum bicolor* and maintained in a growth chamber. It was purified from the symptomatic sorghum leaves according to Masumi *et al.* 2000. The viral RNA was extracted from purified virus using a mRNA Capture Kit (Roche, Germany) according to manufacturer's instruction. Reverse transcription was performed using M-MuLV reverse transcriptase (Fermentas, Lithuania) and the N1T reverse primer (5′-GACCACGCGTATCGATGTCGAC(T)17-3′).
The complete genome sequence of MDMV-Ir was obtained by polymerase chain reaction (PCR) using 12 pairs of overlapping specific primers designed from the sequence of the Bulgarian isolate of MDMV (MDMV-Bg, GenBank Acc. No. AJ001691) and a pair of degenerate primers for potyviruses to amplify the 3′-region of the genome. The 5′- terminus was amplified by sp2 and sp3 primers (Table 1) and one forward primer (PotF0) corresponding to the first 20 nucleotides conserved in potyviruses, designed from MDMV-Bg (Achon *et al.* 2007) (Table 1). PCR consisted of 35 cycles of 94°C for 1 min, 72°C for 1 min and an annealing temperature according to the primer pair used (Table1). Initial denaturing was at 94° C for 4 min and final extension at 72° C for 15 min. The PCR products were ligated into the pTZ57R/T plasmid vector and cloned in *E. coli* strain XL-Blue using InsT/A clone PCR Product Cloning Kit (Fermentas) according to the manufacturer's instruction or electroporation method using multiporator system (Hanahan *et al.* 1991). Recombinant plasmids were extracted using High Pure Plasmid Isolation Kit (Fermentas) and automatically sequenced (Tech Dragon, Hong Kong). Some PCR products were directly sequenced using specific primers for each segment.

Table1. Features of the primers designed for MDMV genome sequence by PCR.

Primer pair	5' → 3' Sequence	Direction	PCR product length (bp)	Position of segment	Annealing temperature
Pot F0 / SP3	5′-aaaaacaacaaractcaacacaacacaac -3′	Forward	237	1-237	55
	5′-ggagctgttcgctgcaaagg -3′	Reverse			
Pot F0 / SP2	5′-aaaaacaacaaractcaacacaacacaac -3′	Forward	196	1-196	55
	5′-cagccaaaccttgacaaccc -3′	Reverse			
MDM1F /	5′-ccttgacaaccctcgtgacg-3′	Forward	603	184-787	55.5
MDM1R	5′-gtggcaagctatggcgttat -3′	Reverse			
MDM2F /	5′-acacaaagggcacagtggtc -3′	Forward	1140	712-1852	53
MDM2R	5′-cacagaccctccaacgatgt -3′	Reverse			
MDM3F /	5′-acagagtttggacaacctg -3′	Forward	1158	1731-2889	55.5
MDM3R	5′-gtcgcagtcgagtcagctct -3′	Reverse			
MDM4F /	5′-tctccgaaacatggcgtgtg -3′	Forward	1198	2737-3935	55.5
MDM4R	5′-gacatcaggctttgcacttca -3′	Reverse			
MDM5F /	5′-agagccgacacgaccactcg -3′	Forward	1188	3791-4979	54
MDM5R	5′-gaggcattcctgagaaagtgtattc -3′	reverse			
MDM6F /	5′-tgatgccatttgagctttcacc -3′	Forward	1259	4717-5976	54
MDM6R	5′-ctccacacgaaccactaagg -3′	reverse			
MDM7F /	5′-tcacagcaattcgagaagctg -3′	Forward	1148	5828-6976	54.5
MDM7R	5′-cacgcacgaggaagctgaga -3′	reverse			
MDM8F /	5′-tccaacatgaaacatggctaacg -3′	Forward	1132	6829-7961	55.5
MDM8R	5′-cacttcgcagcgttagggtt -3′	reverse			
MD1F /	5′-gatgaattgaatgtctatgcacga -3′	Forward	1277	8222-9498	57
MDMVG-SH	5′- gagtttcgtggtgagag-3′	reverse			
Pot2 / Pot1*	5′-gacgaattctgtgatgctgatggtttc -3′	Forward	1192	7560-8752	53
	5′-gactggatccattttctatgcaaca -3′	reverse			

*- The primer pairs designed by Colinet and Kummert, 1993

Genome analyses

MDMV-Ir complete sequence was aligned with 25 complete sequences of cereal potyviruses from GenBank and five complete sequences of MDMV isolates. The CP region of MDMV was also aligned with those of 23 isolates from GenBank representing MDMV isolates of the world (Table 4). Sequence analyses and comparisons were performed using the Clustal X program (Thompson *et al.* 1997) and MEGA5 (Tamura *et al.* 2011). Phylogenetic analysis was conducted by neighbor- joining (NJ) and maximum-likelihood (ML) methods and trees were reconstructed by bootstrapping robustness with 100 replicates to evaluate the significance of the internal branches. The tree was visualized in MEGA5 package or by using TreeView 1.6.6 program (Page, 1996). Pairwise distance between sequences or within phylogenetic groups and the diversity index (π) were calculated by MEGA5 (Tamura *et al.* 2011).

The CP sequences were scanned using the SimPlot (Ray 1999), Bootscan (Salminen *et al.* 1995) and MaxChi (Posada and Crandall 2001) methods implemented in RDP3 program, for possible recombination events. The window and step sizes were adjusted to 200 nucleotides and 20 nucleotides, respectively. Sequence comparisons were performed without correction, based on default adjustment.

Evolution parameters

The nucleotide substitution model was estimated by MEGA5 for further analysis. The ratio (ω) of non-synonymous (amino acid change) and synonymous (silent) substitutions which provide an estimate of the selective pressure at the protein level (Kimura 1983) was estimated using ML for the CP sequence of the 23 selected isolates of MDMV.

To detect positive selection at single amino acid sites, we estimated the rates of non-synonymous and synonymous substitution changes at each site in a sequence alignment using ML- based methods as implemented in MEGA 5 (Tamura *et al.* 2011).

To understand the ancestral relationships and the direction of the evolution of extants via

Table 2. Cleavage sites of polyprotein of MDMV-Ir and some other cereal potyviruses

Potyvirus species	Accession number	P1/HC-Pro	HC-Pro/P3	P3/6K1	6K1/CI	CI/6K2	6K2/VPg	VPg/NIa	NIa/NIb	NIb/CP
JGMV	Z26920	KQICHY/S	KEYIVG/G	TEVEHE/R	QEVKHE/G	QTVIHE/N	TEVEHE/G	PEVEHE/G	ERISNE/S	VDVEHQ/S
MDMV-Bg	AJ001691	QEIEHY/A	REYAVG/G	VGVIHE/G	SQVTHQ/S	VTVIHQ/G	TDVKHE/A	VEVEHE/A	FDVTEQ/G	IDVKHQ/A
MDMV-Sp	AM110758	QEIEHY/A	REYAVG/G	VGVIHE/G	TQVTHQ/S	VTVIHQ/G	TDVKFE/G	VEVEHE/A	FDVTEQ/G	IDVKHQ/A
MDMV-Ir	JO280313	QEIEHY/A	REYAVG/G	VGVIHE/G	TQVTHQ/S	VTVIHQ/G	TDVKHE/A	VEVEHE/A	FDVTEQ/G	IDVKHQ/A
MDMV-Italy	JX185302	QEIEHY/A	REYAVG/G	VGVIHE/G	TQVTHQ/S	VTVIHQ/G	TDVKHE/A	VEVEHE/A	FDVTEQ/G	IDVKHQ/A
MDMV-USA	JQ403608	QEIEHY/A	REYAVG/G	VGVIHE/G	TQVTHQ/S	VTVFHQ/G	TDVKHE/A	VEVEHE/A	FDVTEQ/G	IDVKHQ/A
MDMV-USA	JQ403609	QEIEHY/A	REYAVG/G	VGVIHE/G	TQVTHQ/S	VTVIHQ/G	TDVKHE/A	VEVEHE/A	FDVTEQ/G	IDVKHQ/A
PenMV	DQ977725	LDIDHY/A	RDYIVG/G	TGVIHE/H	KNVVHQ/S	NTVIHQ/G	QDVTHQ/G	EGVTHE/A	DDVMEQ/G	EDVYHQ/S
SCMV	EU091075	MEIEHY/A	REYIVG/G	TGVIHE/G	PPVTQQ/S	NTVIHQ/G	TEVSHQ/G	T(A)GVAHE/S	T(M)SVEEQ/C	EDVFHQ/S
SCMV	AJ310105	F(L)DIEHY/A	REYIVG/G	TGVIHE/G	PPVTQQ/S	NTVIHQ/G	TNVSHQ/G	T(A)GVAHE/S	M(I)SVEEQ/C	EDVFHQ/S
SrMV	U57358	NEITHF/S	REYVVG/G	TGVIHE/A	SLLGTQ/V	TTVIHQ/G	TNVCHQ/G	VGVEHE/A	CEVTEQ/G	IDVRHQ/A

analysis of available nucleotide sequences of CP-UTR region of MDMV isolates and complete nucleotide sequences of other cereal potyviruses, attempts were made to estimate the ancestral (extinct) sequences of the viruses by using MEGA5 (Tamura *et al.* 2011). All extant sequences of MDMV isolates and the estimated ancestral sequences (nodes) were analyzed together.

Results

Twelve overlapping contigs covering the full length of the genome of the MDMV-Ir isolate from Golestan province were obtained and the complete genome sequence was reconstructed and deposited in EMBL database under the GenBank accession number JO280313. It contained 9499 nucleotides, excluding the 3′ poly (A) tail, consisted of one ORF starting with ATG at nucleotide position 139 and terminating by a stop codon TGA at nucleotide position 9262. The ORF encodes a polyprotein of 3042 amino acids (aa), as is typical for other potyviruses. The base composition of the genome of MDMV-Ir isolate was adenine %34.0, cytosine %19.51, guanine %21.13 and uracil %25.33 which is similar to that of other potyviruses. The highly conserved G1–2A6-7 motif in the P3 gene involved in the expression of the ORF PIPO (Chung *et al.* 2008) is also detected as a result of (+2) frame-shift from nucleotide 2678 to 2685 nucleotides in the P3 gene of the MDMV-Ir. The reading frame of the motif in MDMV-Ir is TGAA AAA AA (spaces separate polyprotein codons).

Multiple sequence alignment of MDMV-Ir and three European MDMV isolates (Acc. Nos. JX185302, AJ001691 and AM110758) showed identical size of each predicted gene as well as 5′-UTR and 3′-UTR regions. In comparison, two US isolates (Acc. Nos. JQ403608 and JQ403609) had a 57 nucleotide gap (nucleotide 8470 to 8526) corresponding to 19 amino acids in the coat protein. No gap was found in a third US isolate (A 34974), although, the latter was similar to other US isolates in total CP gene sequence.

The 5'-UTR region with 139 nucleotides shares 85.6 -94.2% identity with other MDMV isolates and 49.7-73.1% identity with other cereal potyviruses. The 5′-UTR begins with a conserved 13 nucleotide stretch, AACACAACACAAC, the so-called potybox (a), which has a role in enhancing translation (Kong and Steinbiss 1998). The 3′-UTR is 234 nucleotides long, excluding the terminal poly (A) tail.

Comparison of the MDMV polyprotein with those of other potyviruses revealed nine potential cleavage sites (Table 2). All MDMV isolates have the same cleavage sites with two exceptions. In 6k2/VPg cleavage site of MDMV-Sp, histidine and alanine residues are replaced by phenylalanine and glycine, respectively. Also in the CI/6K2 cleavage site of the US isolate JQ403609, isoleucine is substituted by phenylalanine. The MDMV-Ir genome contained no deletions or insertions relative to European MDMV isolates.

Specific motifs described for potyvirus polyproteins (Atreya *et al.* 1990; Berger *et al.* 2005; Dougherty and Semler 1983; Hema *et al.* 1999; Shukla *et al.* 1991) were all present in the polyprotein of MDMV-Ir, and almost all were highly conserved (Table 3). Pairwise comparison of the complete genome and polyprotein and the 5′- and 3′- untranslated regions or individual genes of MDMV-Ir with those of other isolates and species of SCMV-subgroup are shown in Table 4. The comparative analyses of the

Table 3. Position of motifs found in different virus genes of MDMV

CP	NIb	CI	P3	HC-Pro	P1
DAG MVWCIENG	NGDDL GNNS GDD	GAVGSGKST	$G_{1\text{-}2}A_{6\text{-}7}$	FRNK PTK	$H_{153}(8)D_{162}\_(31)G_{194}\_S_{196}$ $F_{207}IVRGR_{212}$

sequences revealed that MDMV-Ir shares 93.9% to 97.3% similarity at amino acid level with other MDMV isolates and 91.9 and 86 percent identities at nucleotide level with Bulgarian and US isolates, respectively. Among other potyviruses, MDMV-Ir was closest to Texas isolate of SrMV (U57358) with 77.5% amino acid similarity, and to Yuhang isolate of SrMV from China (AJ310198) with 65.1% nucleotide identity. It had the lowest identity of 50.2% and 49.2% at nucleotide and amino acid levels, respectively, with JGMV isolate (Table 4).

Among various genes of sequenced cereal potyviruses, MDMV-Ir shared the greatest sequence identity of 88% in the CI with SrMV (NC-004035).

The correlation of nucleotide identities of individual genes with the complete nucleotide sequence identity or polyprotein similarity was analyzed by t-test. There was a high correlation, often more than 0.9 between nucleotide identity values pertaining to individual genes and the complete genome. There was no significant difference between the two analyses in t-test (Table 5).

Phylogenetic relationships among SCMV-subgroup viruses were also analyzed using the nucleotide sequence of each of the twelve genome regions. The topology of both ML and NJ trees were essentially similar. Phylogenetic analyses of the complete nucleotide sequence of MDMV-Ir with 22 reported full length sequences revealed that MDMV-Ir is located on a branch with five other isolates of MDMV. SrMV isolates formed a

Table 4. Percentage similarities of amino acid sequences of coding regions and nucleotide sequences of noncoding regions between MDMV-Ir and other cereal potyviruses

Potyvirus Species	Accession number	Complete genome (nucleotide)	Polyprotein (amino acid)	5'UTR	P1	HC-Pro	P3	6K1	CI	6K2	VPg	NIa	NIb	CP	3'UTR
JGMV	NC-003606	50.2	49.2	53.8	29.8	46.2	32.9	32.8	55.6	39.6	56.6	43.4	59.5	61	19.5
JGMV	Z26920	50.2	49.2	53.8	29.8	46.4	32.9	32.8	55.6	39.6	56.6	43.4	59.5	61	19.5
MDMV	AJ001691	91.9	97.3	94.2	89.5	98	96.5	98.5	99.4	94.3	98.4	97.1	98.1	97.1	96.7
MDMV	AM110758	86.4	93.9	93	85.1	96.5	91.6	95.5	98.3	92.5	87.8	95.5	92.1	96.4	96
MDMV	JX185302	86.7	95.4	88.6	85	98	91.1	98.5	99.7	98.1	93.1	96.7	94.2	96.2	92
MDMV	JQ403608	86.6	94.7	88.5	85.8	97.4	90.2	98.5	98.6	94.3	95.2	97.1	95	91.1	92.4
MDMV	JQ403609	86	94.4	85.6	86.7	96.5	88.8	98.5	97.6	94.3	95.2	97.5	95.4	91.4	92.4
PenMV	DQ977725	62.9	71.5	49.7	33.9	79	64.8	65.7	80.1	66	66.7	69.4	77	76.6	72.2
PenMV	AY642590	63.2	71.6	51.5	34.3	79.2	65.4	65.7	80.5	66	66.1	70.2	77	76.9	72.4
PenMV	NC-007147	63.2	71.6	51.5	34.3	79.2	65.4	65.7	80.5	66	66.1	70.2	77	76.9	72.4
SCMV	EU091075	63.1	74.4	71.9	52.4	84.4	68.3	71.6	81.4	64.2	76.2	69	75.2	73.1	73.4
SCMV	AY569692	63.4	74.7	72.5	53.2	83.9	68.6	71.6	81.7	66	75.1	70.2	76	73.1	71.4
SCMV	GU474635	62.9	73.8	73.1	52.8	83.7	68.6	71.6	81.9	67.9	75.7	68.6	75.2	67.6	73.4
SCMV	AY149118	63.2	74.8	72.5	54	83.9	68.9	71.6	82	67.9	74.1	69.4	76.4	72.8	70.5
SCMV	AY042184	63	74.7	71.9	53.6	84.2	68.6	70.1	81.7	67.9	73.5	69.4	76.4	73.1	71
SCMV	AF494510	63.1	75	71.9	53.2	83.9	70.9	70.1	81.9	67.9	73.5	70.2	76.4	73.1	70.7
SCMV	NC-003398	62.8	74.4	71.9	53.2	83.1	68.6	70.1	81.6	67.9	73.5	69.8	76.2	73.1	71.2
SCMV	AJ310105	63.2	73.9	67.8	46.8	83.9	67.7	73.1	81.6	58.5	75.1	70.2	75.2	73.7	71.2
SCMV	AJ310103	63.2	74.2	68.4	50.8	83.3	68.3	70.1	81.9	60.4	76.7	69.8	76	72	75.4
SCMV	AJ310102	63.2	74	69	50	83.5	68.9	70.1	81.4	60.4	77.2	70.2	75.6	70.2	75.4
SCMV	AJ297628	62.8	74.4	71.9	53.2	83.1	68.6	70.1	81.6	67.9	73.5	69.8	76.2	73.1	71.2
SrMV	U57358	64.3	77.5	58.8	53.2	85.9	72	60.3	87.5	73.6	77.8	75.2	78.7	72.8	74.9
SrMV	NC-004035	64.9	76.5	59.6	48.4	82	71.2	62.7	88	71.7	77.8	71.9	78.5	77.2	77.6
SrMV	AJ310198	65.1	76.9	60.2	50	84.2	71.2	62.7	87.8	73.6	77.8	72.7	77.9	77.5	77.6
SrMV	AJ310197	64.9	76.5	59.6	48.8	82	71.2	62.7	88	71.7	77.8	71.9	78.5	77.2	77.6

Table 5. The correlation (r values) of complete genome sequence identities and polyprotein similarities of cereal potyviruses implemented in Table 4 with nucleotide identities and amino acid similarities of individual genes and untranslated regions

	Complete genome	Polyprotein
5'UTR	0.725 *	0.713
P1	0.943	0.918
HC-Pro	0.815	0.949
P3	0.918	0.993
6K1	0.929	0.964
CI	0.892	0.981
6K2	0.952	0.975
VPg	0.939	0.96
NIa	0.97	0.992
NIb	0.98	0.986
CP	0.973	0.929
3'UTR	0.817	0.943

*. There was no significant difference between r values of the two groups. The p-values in all pairwise comparisons were 0.0.

sister clade close to MDMV clade (Fig 1). Phylogenetic trees obtained by analyses of individual genes were topologically congruent with the tree obtained by analyses of the complete genomes (data not shown).

Phylogenetic analysis was also performed on CP-UTR using 24 MDMV sequences deposited in GenBank. The branches of NJ tree constructed (Fig 2) were not supported by high bootstrap values. Therefore, the tree was condensed and a majority – rule bootstrap cutoff of 50% was applied to construct the consensus tree. The resulting tree represented a polytomy, excluding two Chinese divergent isolates (Acc. Nos. S77088 and AY660663). Pairwise nucleotide sequence diversity within MDMV population ranged from 0.0% between some European isolates to 25% between China isolates and others. The Chinese isolates are the most divergent isolates and form a distinct clade in phylogenetic tree. Although the phylogenetic tree consists of a largely unstructured assemblage, it is not completely devoid of structure. It could be further divided into two clades designated group I with ten and group II with nine isolates. Isolates of each group shared a common node of significant bootstrap support (Fig 2). Group I included two isolates from northern Iran and eight isolates from Hungary and Bulgaria. Pair wise nucleotide diversity within group I ranged between 4.2% and 7.3% with π= 0.082. A 518 nucleotide sequence (GenBank Acc. No. EV109535) of another MDMV isolate from northern Iran is also clustered in this Group.

Group II comprised three isolates from Hungary, one from Italy, three from USA, one from Argentina and one from Spain. Pair wise nucleotide sequence diversity within this group ranged from 5.4% to 7.9% with π =0.077. The polymorphism estimated for all MDMV isolates based on θ (the summary statistics of the number of polymorphic sites) was 0.1067 and π = 0.129. The non-synonymous substitution (dN), the synonymous substitution (dS), and the ω = dN/dS ratio calculated for all isolates of MDMV in phylogenetic analysis were 4.4094, 0.8825 and 0.02, respectively.

The Gorgan isolate was the closest isolate to MDMV-Ir, while two other isolates previously reported from Iran, Sari (EU 109534) and Isfahan (EU109533) with 88.8% and 87% nucleotide identity in CP-UTR region with MDMV-Ir were placed in other branch of the tree.

In analyses of CP- region among cereal potyviruses, several putative recombination events were identified among the sequences examined but found to be not statistically significant. There was, therefore, no evidence of recombination in these viruses (data not showed).

In ancestral analyses, the common ancestor of the group which formed by MDMV-Ir, Bulgarian (AJ001691 and AM491607) and Hungarian (FM883228 and FM883181) was placed close to MDMV-Ir, while ancestor of branch which is formed with two other Iranian isolates (EU109533 and EU109534) and a Hungarian isolate (FM883167) was placed close to FM883167. In group II, the USA isolates (A34974, JQ403608 and JQ403608) and AM110758 from Spain share a subclade and DQ973169 from Argentina forms another subclade of Hungarian isolate (Fig 3).

Discussion

Comparison of MDMV-Ir with other species of SCMV-subgroup shows that it is closely related to European MDMV isolates. This may suggest that the Iranian isolate has originated from Europe. However, a related virus, BgSMV, is endemic in southern Iran and widely distributed

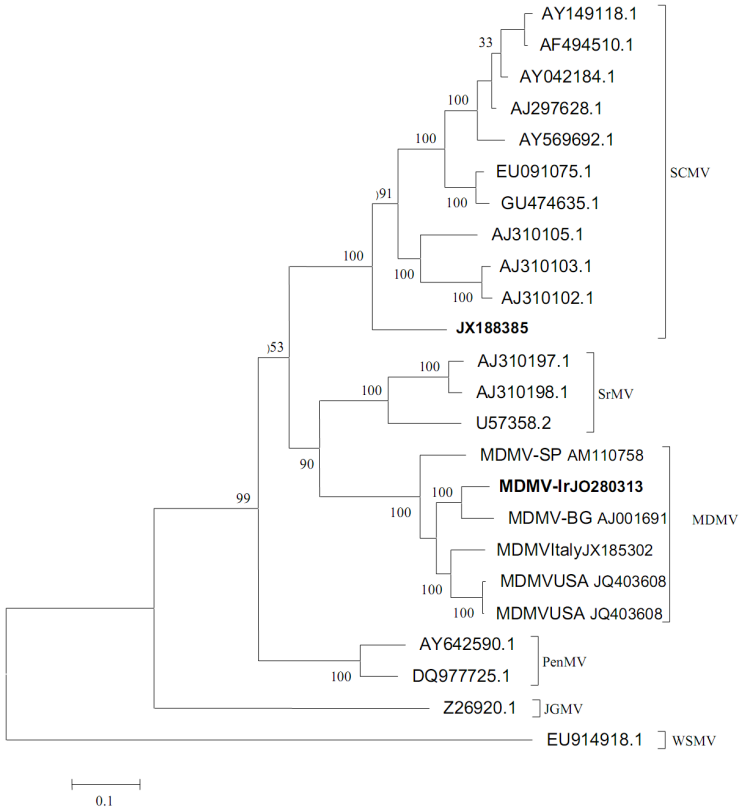

Fig.1. Maximum-likelihood analysis based on alignment of nucleotide sequence of complete genome of MDMV-Ir and cereal potyviruses. Wheat streak mosaic virus (WSMV) was used as out-group.

on Bermuda grass. BgSMV is the closest virus to MDMV (Zare *et al.* 2006; Farahbakhsh *et al.* 2013; Zakeri *et al.* 2012). It is probable that MDMV and BgSMV have diverged from each other. Among other cereal potyviruses, ancestor of MDMV and SrMV is closer to MDMV than to SrMV. SrMV has been reported from the USA and Australia but its plant of origin has remained unknown. MDMV and IJMV may have originated in Johnson grass in the Mediterranean basin. However, the most recent common ancestor of cereal potyviruses (MDMV, SCMV, SrMV, and IJMV) has more proximity to IJMV than to other viruses. It shows that probably both MDMV and IJMV have evolved in Johnson grass. Johnson grass is native to Mediterranean region and grows throughout Europe and the Middle East. It is believed to have arrived in the United States in the early 1800s as a potential forage crop or mixed in flax seed harvested in the Middle East in 1930's (Holm et al. 1977). Thus, based on host-virus coevolution hypothesis (Fraile and Garcia-Arenal 2010), MDMV and Johnson grass appear to have the same epicenter in the Mediterranean.

The topology in the phylogram depicted based on CP-UTR analyses (Fig 2) indicate that the populations of MDMV are not well structured due to rapid distribution in the Mediterranean region where most populations of the virus were examined. Although MDMV was first reported from USA (Janson and Ellett 1963), most sequences deposited in GenBank are from Europe. It seems that most isolates of MDMV are derived from few isolates as the common ancestors but that multiple lineages have subsequently diverged in a radial fashion via seed or vector transmission. The original host of MDMV seems to be Johnson grass from which the virus has been transferred to maize as a result of extensive cultivation of this crop in Europe and Asia. High pairwise sequence identity values among MDMV isolates and low number of polymorphic sites have been taken as evidence of low genetic diversity and high genetic stability (Garcia-Arenal *et al.* 2001). Likewise, the low ω value obtained suggests a strong action of purifying selection in the evolution and variation

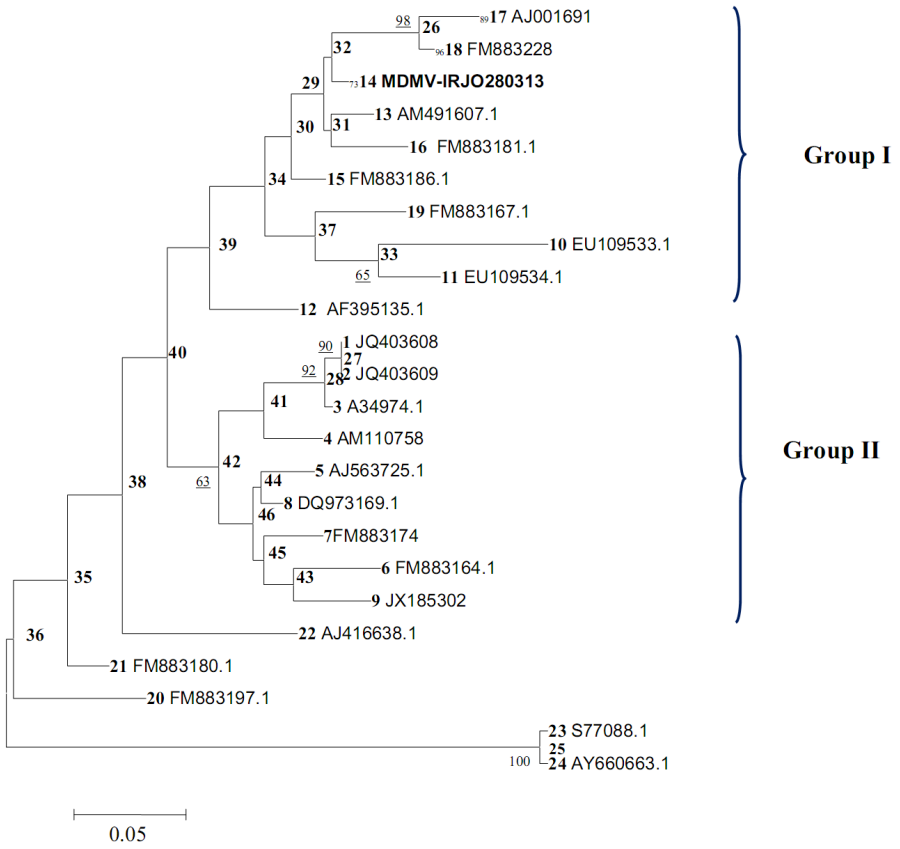

Fig.2. Maximum-likelihood analysis based on alignment of CP-UTR nucleotide sequence of MDMV representative world isolates. The isolates are numbered (1-24) and shown as accession numbers. Node numbers are bolded. Bootstrap values are underlined. Branches with less than 50% bootstrap support have been omitted. The isolates of MDMV used in this study, their accession number and geographical origin were as follows: MDMV-Bg (AJ001691) from Bulgaria, MDMV- Sz0710 (FM883228), Sz0801 (FM883181), Sz0806 (FM883186), Sz0804 (FM883168), Mv0817 (FM883180), Mv0814 (FM883197), Dallas/H (AJ563725) from Hungary; MDMV- Isfahan (EU109533) and MDMV- Sari (EU109534) from Iran; MDMV- Israel (AF395135) from Israel; MDMV- OH-1 (JQ403608) and MDMV- OH-2 (JQ403609) from the USA; MDMV-A (A34974); MDMV-Arg (DQ973169) from Argentina; MDMV- Italy (JX185302) from Italy; MDMV-SP (AM110758) and MDMV-S2 (AJ416638) from Spain; MDMV (S77088) and MDMV-Hebei (AY660663) from China.

of MDMV populations in the world. The homology of the CP sequences suggests that either due to frequent gene flow the time available for individualization has been short, or MDMV CP has not been subjected to great evolutionary pressure.

Ancestral analysis (Fig 3) shows that the four isolates from American continent have been probably originated from Europe, or specifically from the Mediterranean basin. The node of subclade containing the USA and Spanish isolates is closer to Spanish than to USA isolates. In spite of the polytomy structure of the group II subclade in phylogenetic tree and ancestral tree, it seems that the most recent common ancestor of this group is originally from Europe. This clade is a monophyletic group and high sequence identity of USA and Spain sequences suggests that MDMV has been introduced to the American continent via weed seeds. The number and direction of introductions cannot be ascertained. Based on these analyses, it seems that, the most recent common ancestor of the MDMV isolates in the world has been developed in the Mediterranean basin, and then disseminated in other areas and continents.

The percent identities calculated for complete genome sequences had a high correlation with

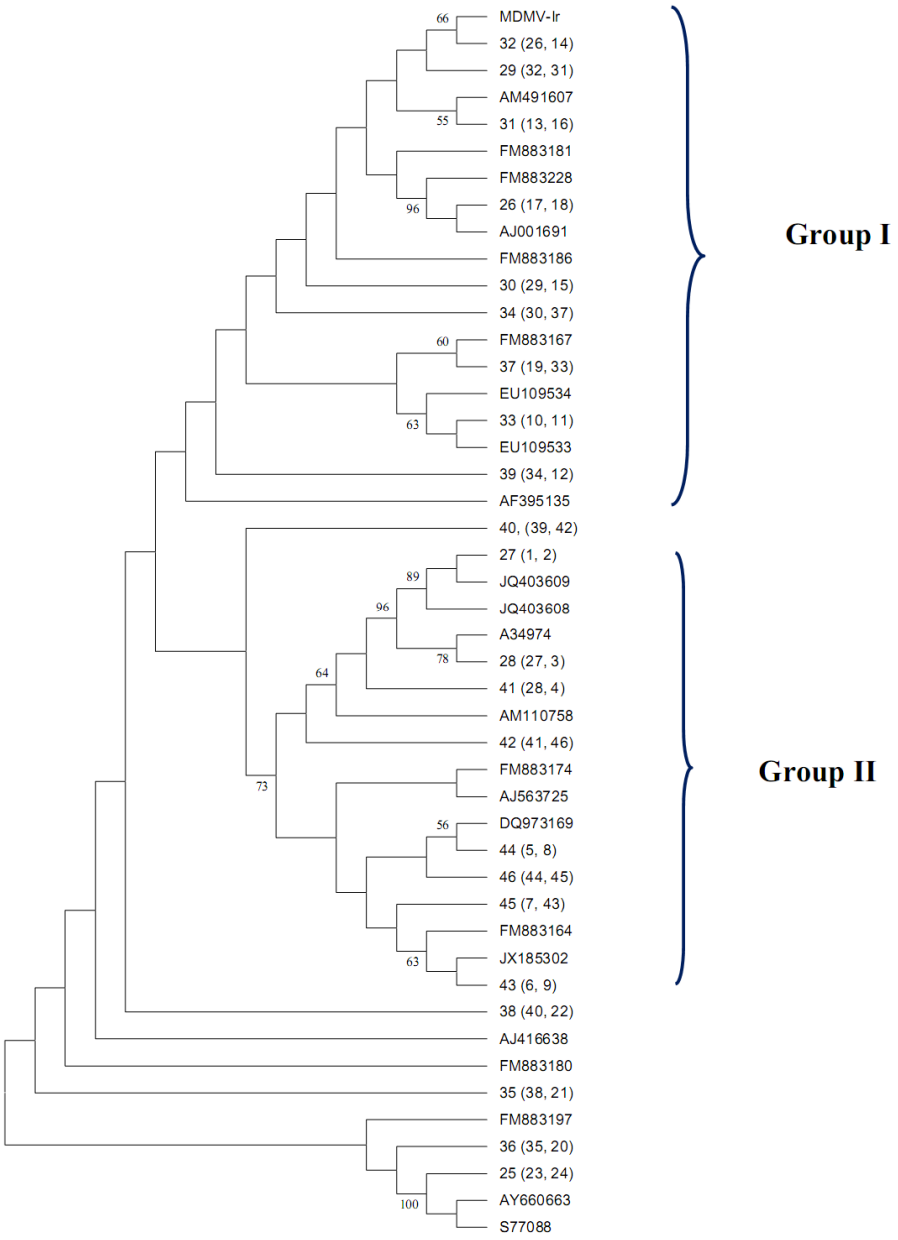

Fig. 3. Phylogenetic relationship of CP-UTR among MDMV isolates of Fig. 2 with their ancestral analyses based on maximum likelihood using MEGA5. The numbers in parentheses are estimated pairs of extant sequences related to each ancestor node. The numbers refer to nodes of Fig. 2. Branches with less than 50% bootstrap support have been omitted. See legend of Fig. 2 for details of accession number and geographical origin

identity values of individual genes, suggesting that the relationship based on each gene would be an estimate of overall relatedness of viruses and any of the genes or untranslated regions of potyviruses can be used in phylogenetic analyses. Although Adams *et al.* (2005) suggested that CI is the best representative of complete genome for evolution and taxonomic studies in *Potyviridae*, we found no significant difference between correlation of CI identity value (r= 0.827) with that of CP identity value (r= 0.973) and identity value of the complete genomes (Table 5). The highest nucleotide identity among the species is found in the CI gene. This indicates that the CI is

the most conserved and P1 the most variable gene in the genome of potyviruses. However, the CP gene is often used to study the diversity and evolution of potyviruses. The CP gene is also responsible for serological relationships, symptom expression, and vector relationship (Urcuqui-Inchima et al. 2001). Because of its higher relevance to the biological properties of the viruses, it may be a better representative of the whole genome than other genes.

Acknowledgements

The authors wish to thank A. Rowhani of the University of California at Davis, CA, for reviewing the manuscript prior to submission. This research was supported in part by the Iranian Council of Centers of Excellence.

References

Achon, M.A., Serrano, L., Alanso-Duenas, N. and Portan, C. 2007. Complete genome sequences of Maize dwarf mosaic and Sugarcane mosaic virus isolates coinfecting maize in Spain. Arch. Virol. 152: 2073–2078.

Adams, M.J., Antoni, J.F. and Fauquet, C.M. 2005. Molecular criteria for genus and species discrimination within the family Potyviridae. Arch. Virol. 150: 459-479.

Atreya, C.D., Raccach, B. and Pirone, T.P., 1990. A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. Virology 178:161-165.

Berger, P.H., Adams, M.J., Barnett, O.W., Brunt, A.A., Hammond, J., Hill, J.H., Jordan, R.L., Kashiwazaki, S., Rybcki, E., Spence, N., Stenger, D.C., Ohki, S.T., Uyeda, I., van Zaayen, A., Valkonen, J., and Vetten, H.J. 2005. Family Potyviridae.Pp. 819-841, *In* C.M. Fauquet, M.A. Mayo, J. Maniloff, U. Desselberger and L.A. Ball (eds.). Virus Taxonomy, Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA.

Chung, B.Y-W., Miller, W.A., Atkins, J.F. and Firth, A.E. 2008. An overlapping essential gene in the Potyviridae. Proc Natl Acad Sci USA 105: 5897–5902.

Colinet, D. and Kummert, J. 1993. Identification of a sweet potato feathery mottle virus isolate from China (SPFMV-CH) by the polymerase chain reaction with degenerate primers. J. Virol. Methods 45:149-159.

Converse, R.H. and Martin, R.R. 1990. ELISA methods for plant viruses. Pp. 179-196, *In* R.O. Hampton, E.M. Ball, S.H. DeBoer (eds.). Serological Methods for Detection and Identification of Viral and Bacterial Plant pathogens, A Laboratory Manual. APS Press, USA.

Dougherty, W.G. and Semler, B.L. 1993. Expression of virus-encoded proteinases: functional and structure similarities with cellular enzymes. Microbiol. Rev 57: 781–822.

Fan, Z., Chen, H., Cai, S., Dong, C., Wang, W., Liang, X. and Li, H. 2003. Molecular characterization of a distinct potyvirus from whitegrass in China. Arch. Virol. 148: 1219-1224.

Fan, Z. and Li, H. 2004. Pennisetum mosaic. Pp. 687-689, *In*: H. Lapierre, P-A Signoret (eds.). Viruses and Virus Diseases of Paoceae (Graminae). INRA Edition Paris, France.

Farahbakhsh, F., Masumi, M., Afsharifar, A., Izadpanah, K. and Rahpeyma-Sarvestani, N. 2013. Phylogenetic analysis of Bermuda grass southern mosaic virus isolates in Iran. Iran J Plant Path. 49: 61-75.

Frail, A. and Gracia-Arenal, F. 2010. The coevolution of plants and viruses: resistance and pathogenicity. Adv. Virus Res. 76:1-32.

Gracia-Arenal, F., Frail, A. and Malpica, J.M. 2001. Variability and genetic structure of plant virus populations. Annu. Rev. Phytopathol. 39: 157-186.

Hanahan, D., Jessee, J.R. and Bloom, F. 1991. Plasmid transformation of *Escherichia coli* and other Bacteria. Abs. Meth. Enzymology 24: 63.

Heydari, S., Izadpanah, K., Masumi, M. and Afsharifar, A. R. 2008. Phyogenetic analysis of CP-UTR in the genome of Reed canary grass mosaic virus. Proc.18th Iranian Plant Protec. Cong., Hamedan, Iran 2008: 475.

Hema, M., Joseph, J., Gopinath, K., Sreenivasulu, P. and Savithri, H.S. 1999. Molecular characterization and interviral relationships of a flexuous filamentous virus causing mosaic disease of sugarcane (*Saccharum officinarum* L.) India. Arch. Virol. 144: 479–490.

Hollongs, M.L. and Brunt, A.A. 1981. Pp. 731-807 Potyviruses, *In* E Kurstak (eds.). Handbook of plant virus infection comparative diagnosis. Elsevier, North Holland.

Holm, L.G., Donald, P., Pancho, J.V. and Herberger, J.P. 1977. The World's Worst Weeds: Distribution and Biology. The University Press of Hawaii, Honolulu, Hawaii, USA.

Janson, B.F. and Ellett, C.W. 1963. A new corn disease in Ohio. Plant Dis. Rep. 47: 1107-1108.

Kimura, M. 1983. The Neutral Theory of Molecular Evolution. Cambridge: Cambridge University Press, UK.

Kong, P. and Steinbiss, H.H. 1998. Complete nucleotide sequence and analysis of the putative polyprotein of maize dwarf mosaic virus genomic RNA (Bulgarian isolate). Arch. Virol. 143: 1791–1799.

Masumi, M. and Izadpanah, K. 1998. Bermudagrass mosaic in Iran. Proc. 13th Iran. Plant Protec. Cong., Karaj, Iran. 314

Masumi, M., Izadpanah, K. and Lesman, E. 2000. Purification and serology of wheat eqlid mosaic virus. Proc. 14th Iran. Plant Protec. Cong., Isfahan, Iran. 20. (Abst.).

Masumi, M., Izadpanah, K. and Behjatnia, S.A.A. 2001. Taxonomical position of Iranian Johnson grass mosaic virus. Proc. 1st Iran. Virol. Cong., Tehran, Iran. Pp. 136-137. (Abst.).

Masumi, M., Zare, A. and Izadpanah, K. 2004. Analysis of 5' region of coat protein gene confirms presence of Maize dwarf mosaic virus (MDMV) in widely separated location in Iran. Proc. 2nd Iran. Virol. Cong., Tehran, Iran. Pp.369-370. (Abst.).

Page, R.D.M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12: 357–358.

Posada, D. and Crandall, K.A. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. In Proc. Natl. Acad. Sci. U.S.A. 98:13757-13762.

Ray, S.C. (1999). SimPlot for windows, version 2.5 (Distributed by the author, http//w.w.w.med.jhu.edu/depmed/sry/download/).

Salminen, M.O., Carr, J.K., Burke, D.S. and McCutchan, F.E. 1995. Identifcation of breakpoints in intergenotypic recombinants of HIV type-1 by bootscanning. AIDS Research Human. Retroviruses 11:1423-1425.

Salomon, R. and Seifers, D.L. 2004. Zea mosaic. Pp. 693-696, *In* H Lapierre, P-A Signoret (eds.). Viruses and Virus Diseases of Poaceae (Graminae). INRA Edition Paris, France.

Seifers, D.L., Salomon, R., Marie-Jeanne, V., Alliot, B., Signoret, P., Haber, S., Loboda, A., Ens, W., She, Y.M. and Standing, K.G. 2000. Characterization of a novel potyvirus isolated from maize in Israel. Phytopathology 90: 505-513.

Shukla, D.D. and Ward, C.W. 1988. Amino acid sequence homology of coat proteins as a basis for identification and classification of the potyvirus group. Virus Genes 96: 2703-2710.

Shukla, D.D., Frenkel, M.J. and Ward, C.W. 1991. Structure and function of the potyvirus genome with special reference to the coat protein-coding region. Can.J. Gen. Plant Pathol 13: 178-191.

Stewart, L.R., Bouchard, R., Redinbaugh, M.G. and Meulia, T. 2012. Complete sequence and development of a full-length infectious clone of an Ohio isolate of Maize dwarf mosaic virus (MDMV). Virus Res.165: 219-224.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using Maximum Liklihood Evolutionary Distance and Maximum Parsimony methods. Molecular Biology and Evolution doi: 101093/molbev/msr121.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24: 4876–4882.

Urcuqui-Inchima, S., Haenni, A.L. and Bernardi, F. 2001. Potyvirus proteins: a wealth of functions. Virus Res.74:157-175.

Zakeri, A., Masumi, M., Nasrollanejad, S., Ghahramani, T. and Izadpanah, K. 2012. Cross protection between maize dwarf mosaic virus and Bermuda grass southern mosaic virus. Iran J. Plant Path. 48: 343-352.

Zare, A., Masumi, M., Hayati, J. and Izadpanah, K. 2004. Serological and biological properties of MDMV in Iran.In Proc. 16th Iran. Plant Protec. Cong., Tabriz, Iran 2004: 110.

Zare, A., Masumi, M. and Izadpanah, K. 2005. Bermuda grass southern mosaic virus : A distinct potyvirus infecting several gramineous species in Iran. Parasitica 61: 105-110.

Zare, A., Masumi, M., Heydari, S. and Izadpanah, K. 2006. Report of a new virus associated with mosaic symptom in *Phalaris arundinaceae* in Golpaygan. In: Abbasi M., Aliabadi, F (Eds.). Proc. 17th Iran. Plant Protec. Cong., Karaj, Iran 455 (abst.).