J. Agr. Sci. Tech. (2022) Vol. 24(5): 1233-1249

Genomic Characterization and Phylogenetic Analysis of Two *Potyviruses* Infecting Iris in Iran

A. Naseri[1], Z. Moradi[2], M. Mehrvar[1*], and M. Zakiaghl[1]

ABSTRACT

Several viruses affect iris plants worldwide, and are major constraints in commercial production due to serious economic losses. The first genomic sequences of two *potyviruses*, namely, *Iris Severe Mosaic Virus* (ISMV) and *Iris Mild Mosaic Virus* (IMMV) from naturally infected iris plants (*Iris versicolor*) in Iran were determined using RNA deep sequencing and RT-PCR followed by sequencing of amplicons. Both viruses (ISMV-Ir and IMMV-Ir) had a typical *potyvirus* genetic organization, with a large open reading frame translated as a polyprotein, including nine autocatalytic cleavage sites, and a putative smaller protein P3N-PIPO. Phylogenetic analyses and sequence comparisons revealed close relationships between ISMV and members of group *Onion Yellow Dwarf Virus* (OYDV) of the genus *Potyvirus*. The ISMV-Ir showed > 92% nucleotide (nt) identity (> 96% amino acid (aa) identity) to the three previously reported ISMV isolates, the highest with the Japanese isolate J (94.10% nt identity, 97.41% aa identity) and the lowest with Chinese isolate BJ (92.73% nt identity, 96.77% aa identity). IMMV-Ir belonged to the *Chilli Veinal Mottle Virus* (ChVMV) group of *potyviruses*, had 82.36% nt identity (91.25% aa identity) with the BC32 isolate, and 75.55% nt identity (83.59% aa identity) with the WA-1 isolate from Australia. The genetic distance among IMMV polyprotein-coding genomic sequences or gene-specific sequences indicated a high genetic divergence of these isolates. Our analysis indicated that natural selection has contributed to the evolution of isolates belonging to the two identified *potyviruses*. The information on genomic sequences presented in this study will improve our understanding of virus function and pathogenicity leading to better control of the disease.

Keywords: Genomic sequence, IMMV, ISMV, Natural selection, RNA-Seq.

INTRODUCTION

The family *Potyviridae* is one of the largest and most important groups of plant viruses, being economically significant for a wide range of agricultural crops. *Iris* spp. are perennial bulbous or rhizomatous plants belonging to the family *Iridaceae* (Stevens, 2001), with a cosmopolitan distribution. It is one of the most important and prized groups of plants in horticulture, due to its significant ornamental, cosmetic, and pharmaceutical properties. Viral diseases of iris plants have become widespread, causing serious crop losses around the world. At present, there are several reports of *potyviruses* infecting *Iris* spp. including *Iris Mild Mosaic Virus* (IMMV), *Iris Severe Mosaic Virus* (ISMV), *Iris Fulva Mosaic Virus* (IFMV), *Bean Yellow Mosaic Virus* (BYMV), *Turnip Mosaic Virus* (TuMV), *Ornithogalum Mosaic Virus* (OrMV), *Narcissus Latent Virus* (NLV), *Butterfly Flower Mosaic Virus* (BFMV), and *Gladiolus Mosaic Virus* (tentative name) (Kulshrestha *et al*., 2006a, 2006b; Van der Vlugt *et al*., 1993; Wei *et al*.,

[1] Department of Plant Pathology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Islamic Republic of Iran.

[2] Department of Plant Pathology, Faculty of Crop Sciences, Sari Agricultural Sciences and Natural Resources University, P. O. BOX: 578, Sari, Islamic Republic of Iran.

*Corresponding author; e-mail: mehrvar@um.ac.ir

2006, 2007; Chen *et al*., 2008; Wylie *et al*., 2019). ISMV and IMMV, which cause significant economic damage to iris plants in many countries, can be transmitted in a non-persistent manner by aphids such as *Myzus persicae* and *Macrosiphum euphorbiae*. Viruses of the genus *Potyvirus* have flexuous filamentous particles of about 750×12 nm, with a genome that is composed of a positive-sense single-stranded RNA of approximately 10,000 nt (Ivanov *et al*., 2014; Wylie *et al*., 2017) and a covalently linked Virus-encoded Protein (VPg) is bound to the 5' terminus and a poly (A) tail at the 3′ end (Revers and Garcia 2015). The genomic RNA is translated into a large polyprotein, which after processing by three viral encoded proteases, gives rise to ten mature multifunctional protein products: P1, HC-Pro, P3, 6K1, CI, 6K2, NIa-VPg, NIa-Pro, NIb and CP (Revers and Garcia, 2015; Adams *et al*., 2005). In addition, a short polypeptide (PIPO) is expressed as a result of RNA polymerase transcriptional slippage (Rodamilans *et al*., 2015). In recent years, several isolates of ISMV have been identified and characterized from different countries, including Japan, India, China, Iran, South Korea, Mexico, and Netherlands, based on partial or complete genome sequences (Li *et al*., 2016; Nishikawa *et al*., 2019; Nateqi *et al*., 2015; Kulshrestha *et al*., 2004; Van der Vlugt *et al*., 1994; Yan *et al*., 2010; Park *et al*., 2000). In addition, the partial genome sequences of IMMV have been sequenced in Australia, New Zealand, Netherlands, India and Iran (Wylie *et al*., 2012, 2019; Nateqi *et al*., 2017). Analyses of complete genome sequences of virus isolates provide valuable information for understanding genetic, biological, and epidemiological characteristics, which in turn lead to understanding the evolutionary history and designing new control strategies (Moradi and Mehrvar, 2019). Although the following two *potyviruses* have been reported to infect iris in Iran, however, sequence information for these viruses is only available from the partial of CP and 3'-UTR (for ISMV) and the partial of NIb-CP (for IMMV) and the complete genome sequences have not been determined.

Accordingly, determining the genome sequences of these two important *potyviruses* infecting iris in Iran, where iris is widely grown commercially for cut flowers other than medicinal uses, is necessary. Next-generation sequencing of small RNAs accumulating in plant tissues is increasingly used for identifying previously reported and new viruses and viroids (Seguin *et al*., 2014). In this study, for the first time, the nearly complete genome sequence and organization of ISMV and IMMV from Iran were determined using Illumina RNA-Seq and their genome organization and relationship with other *potyviruses* described.

MATERIALS AND METHODS

RNA Isolation, Library Construction, RNA-Seq and Small RNA Bioinformatics Analysis

In 2019, *Iris versicolor* leaves (n= 7) showing mosaic, mild mosaic, conspicuous chlorotic stripe, and dwarfing symptoms (Figure 1) were collected from urban areas of Mashhad city, Iran. For diagnostic purposes, all the symptomatic leaf samples were pooled for high-throughput paired-end RNA sequencing. Total RNA was extracted from the pooled sample using SV Total RNA Isolation Kit (Promega, USA), followed by the removal of rRNAs using a Ribo-Zero rRNA Removal Kit (Plant Leaf) (Epicentre, USA), both according to the manufacturers' instructions. The library was constructed using a TruSeq RNA Sample Preparation Kit (Illumina, USA) and sequenced using an Illumina NovaSeq 6000 (Macrogen). All reads were trimmed to remove low-quality bases and adaptor sequences and assembled using the *de novo* assembly algorithms of CLC Genomics Workbench v.20 (Qiagen). Assembled contigs were analyzed using BLASTn searches by Geneious Prime software v. 2019.1.3 (Biomatters, New

Figure 1. Characteristic symptoms of severe mosaic and systemic chlorotic stripes on *Iris versicolor* leaves infected with ISMV and IMMV.

Zealand) to identify sequence similarities to reference viral genomes in the GenBank database.

Recombination Analysis

The aligned polyprotein nucleotide sequences were examined for the presence of inter- and intra-lineage recombination events using seven recombination detection methods implemented in the RDP4 software (Martin *et al*., 2015) with a Bonferroni corrected p-value cut-off of 0.01.

Phylogenetic Analysis

Open Reading Frames (ORF), mature peptides, and domains encoded by them were predicted within the NCBI Conserved Domain Database (CDD) and InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan). To investigate the relationship of these viruses to other *potyviruses* (representing 49 species), MUSCLE multiple sequence alignments, and comparison of nucleotide and amino acid sequence identities were carried out using the Geneious Prime. Phylogenetic trees were constructed using the Maximum Likelihood (ML) and Neighbour-Joining (NJ) methods in MEGAX (Kumar *et al*., 2018). ML and NJ trees were built based on complete or nearly complete aa sequences of 56 *potyvirus* isolates from 51 species (Table 1), as well as on 67 nucleotide and aa sequences of the coat Protein Core (cCP) region from the same 51 species, after codon-based alignment with ClustalW (Table 1). The number of Haplotypes (H), Haplotype diversity (Hd), nucleotide diversity (π) value, Nonsynonymous substitution rate (dN) and Synonymous substitution rate (dS) were obtained using the DnaSP6 (Rozas *et al*., 2017).

Biological Assay

The crude sap from symptomatic iris plant leaf (leaf tissue was ground in 0.01M sodium phosphate buffer, pH 7) was mechanically inoculated on carborundum-dusted leaves of *Chenopodium quinoa*, *C. amaranticolor*, *Nicotiana benthamiana*, *N. clevelandii*, *N. glutinosa*, *N. tabacum* cv. *Samsun*, *N. tabacum* cv. *Xanthi*, *Phaseolus vulgaris*, *Vigna unguiculata*, *Cucumis sativus*, *Cucurbita pepo*, *Datura stramonium*, and *Petunia hybrida*. At two to three weeks post-inoculation, the plants were examined for symptoms and assayed for ISMV and IMMV

infection by RT-PCR, using virus-specific primers as described earlier (Li *et al*., 2016; Wylie *et al*., 2019).

Table 1. Dataset of the complete or partial nucleotide sequences of *potyviruses* used in this study.

Virus name (Abbreviation)	Isolate/Strain	Host	Origin	Genome size		Accession number
				nt	aa	
Iris Severe Mosaic Virus (ISMV)	**Ir**	***Iris versicolor***	**Iran**	10454	3316	**MN520154**
	J	*Iris tectorum*	Japan	10403	3316	LC433737
	BJ	Iris	China	10423	3316	KT692938
	Tai'an	Iris	China	10406	3316	MF385582
	PHz*****	Chinese iris	China	1745	461	FJ481099
	TA1*****	*Iris tectorum*	China	753	250	KT998897
	TA2*****	*Iris tectorum*	China	753	250	KT998898
	TA3*****	*Iris tectorum*	China	753	250	KT998899
	TA4*****	*Iris tectorum*	China	753	250	KT998900
	Crocus*****	*Crocus vernus*	Netherlands	2510	723	X75939
	K*****	unknown	Korea	873	290	AF034839
Iris Mild Mosaic Virus (IMMV)	**Ir**	***Iris versicolor***	**Iran**	9159	3052	**MN746770**
	BC32	Iris×hollandica	Australia	9468	3046	MH886513
	WA-1	*Iris* sp.	Australia	8651	2755	JF320812
	Bate2*****	*Iris xiphium*	Australia	6251	2073	JN127338
	DC4a*****	Iris×hollandica	New Zealand	1815	490	DQ436918
	DC4b*****	Iris×hollandica	New Zealand	1813	489	DQ436919
	Blaauw*****	Iris	Netherlands	1430	364	EF203682
Bean Common Mosaic Necrosis Virus (BCMNV)	Michigan (Strain: NL-3)	*Phaseolus vulgaris*	USA	9612	3066	U19287
Bean Common Mosaic Virus (BCMV)	NL1	Common bean	USA	9992	3222	AY112735
Bidens Mottle Virus (BiMoV)	B3	*Bidens* sp.	Taiwan	9709	3071	EU250212
Bidens Mosaic Virus (BiMV)	SP01	*Bidens pilosa*	Brazil	9557	3061	KF649336
Bean Yellow Mosaic Virus (BYMV)	SW9	*Diuris* sp.	Australia	9535	3056	KF632713
Cowpea Aphid-Borne Mosaic Virus (CABMV)	RR3	Cowpea	India	9894	3174	KM597165
Clover Yellow Vein Virus (ClYVV)	Hefei	Broad bean	China	9585	3072	KU922565
Dasheen Mosaic Virus (DsMV)	M13	*Zantedeschia aethiopica*	China	10038	3191	AJ298033
Endive Necrotic Mosaic Virus (ENMV)	FR	Lettuce	France	10000	3236	KU941946
Johnsongrass Mosaic Virus (JGMV)	Brazilian	*Panicum maximum*	Brazil	9874	3058	KT289893
Maize Dwarf Mosaic Virus (MDMV)	OH-1	*Zea mays*	USA	9458	3022	JQ403608
Pennisetum Mosaic Virus (PenMV)	C	*Pennisetum centrasiaticum*	China	9611	3065	DQ977725
Sorghum Mosaic Virus (SrMV)	Xiaoshan	Sugarcane	China	9624	3071	AJ310197
Sugarcane Mosaic Virus (SCMV)	Beijing	*Zea mays*	China	9595	3063	AY042184
Iranian Johnsongrass Mosaic Virus (IJMV)	Maz-Bah	*Zea mays*	Iran	9544	3054	KT899778
Japanese Yam Mosaic Virus (JYMV)	Japanese yam1	Yam	Japan	9757	3130	AB016500
Yam Mosaic Virus (YMV)	YMV-NG	*Dioscorea rotundata*	Nigeria	9594	3103	MG711313
Lettuce Italian Necrotic Virus (LINV)	I234	*Lactuca sativa*	Italy	9829	3173	KP769852
Lily Mottle Virus (LMoV)	Sb	*Lilium* sp. cv. Sorbonne	China	9644	3095	AJ564636
Lettuce Mosaic Virus (LMV)	CL394	*Lactuca sativa*	Chile	10080	3255	KJ161179
Leek Yellow Stripe Virus (LYSV)	Yuhang GYH	Garlic	China	10142	3152	AJ307057
Moroccan Watermelon Mosaic Virus (MWMV)	TN05-76	Zucchini	Tunisia	9730	3124	EF579955
Peanut Mottle Virus (PeMoV)	Para	*Arachis pintoi*	Brazil	9680	3099	MK396065
Pepper Mottle Virus (PepMoV)	128	bell pepper	South Korea	9640	3068	EU586122
Papaya Ringspot Virus (PRSV)	BS10	Pumpkin	China	10327	3344	MK988418
Pea Seed-borne Mosaic Virus (PSbMV)	PSB194CZ	*Pisum sativum*	Czech Republic	9919	3206	MK116871
Potato Virus A (PVA)	B11	*Solanum tuberosum*	Hungary	9585	3059	AJ296311

Table 1 is continued:

Continued Table 1.

Virus name (Abbreviation)	Isolate/Strain	Host	Origin	Genome size		Accession number
				nt	aa	
Potato Virus V (PVV)	AB	*Solanum tuberosum*	Netherlands	9859	3066	MK756119
Potato Virus Y (PVY)	OH	*Solanum tuberosum*	Japan	9699	3061	AB714134
Soybean Mosaic Virus (SMV)	G6H	*Glycine max*	South Korea	9585	3067	FJ640981
Sunflower Chlorotic Mottle Virus (SuCMoV)	Common	*Helianthus annuus*	Argentina	9965	3191	GU181199
Tobacco Etch Virus (TEV)	N	Tobacco	USA	9495	3054	KM282189
Turnip Mosaic Virus (TuMV)	TANX2	Radish	China	9833	3164	EU734433
Vanilla Distortion Mosaic Virus (VDMV)	Cor	*Coriandrum sativum*	India	9553	3083	KF906523
Watermelon Mosaic Virus (WMV)	C07-284	Zucchini	France	10031	3215	JF273468
Zucchini yellow Mosaic Virus (ZYMV)	Kuchyna	*Cucurbita pepo*	Slovakia	9593	3080	DQ124239
Onion Yellow Dwarf Virus (OYDV)	OYDV-At	*Allium cepa*	Germany	10459	3381	JX433020
Narcissus Degeneration Virus (NDV)	Zhangzhou	*Narcissus tazetta*	China	9816	3184	AM182028
Narcissus Yellow Stripe Virus (NYSV)	Marijiniup 3	*Narcissus* sp.	Australia	9647	3106	JQ395042
Vallota Speciosa Virus (VSV)	Marijiniup 7	*Cyrtanthus elatus*	Australia	9908	3102	JQ723475
Shallot Yellow Stripe Virus (SYSV)	ZQ2	*Allium fistulosum*	China	10429	3401	AJ865076
Narcissus Late Season Yellows Virus (NLSYV)	BC37	*Narcissus* sp.	Australia	9665	3105	MH886515
Ornithogalum Mosaic Virus (OrMV)	KP	Diuris	Australia	9445	3015	JQ807997
Beet Mosaic Virus (BtMV)	Wa	*Beta vulgaris*	USA	9591	3085	AY206394
Peace Lily Mosaic Virus (PeLMV)	Haiphong	*Spathiphyllum patinii*	Viet Nam	9882	3079	DQ851494
Chilli Veinal Mottle Virus (ChVMV)	YN-tobacco	Tobacco	China	9739	3089	JX088636
Tobacco Vein Banding Mosaic Virus (TVBMV)	YND	Tobacco	China	9570	3079	EF219408
Daphne Mosaic Virus (DaMV)		Daphne mezereum	Czech Republic	9548	3071	DQ299908
Papaya Leaf Distortion Mosaic Virus (PLDMV)	HaiNan-DF	Carica papaya	China	10153	3269	JX974555

Note: The isolates which used only in cCP relationship analysis are marked with an asterisk. The ISMV and IMMV isolates, which have been sequenced in this study, are bolded.

RRSULTS

Infected iris plants mostly exhibited severe mosaic and systemic chlorotic stripes on leaves, along with dwarfing (Figure 1). However, no symptoms were observed on indicator plants following inoculation of *potyvirus* infected iris leaf sap. These results were confirmed by RT-PCR.

Genome Assembly and Characterization of *ISMV*-Ir and *IMMV*-Ir

The two largest contigs were used to interrogate NCBI databases by BLASTn and BLASTx to identify sequences with shared identity to known viruses. The nearly complete genome sequence of ISMV-Ir (accession number MN520154), and the polyprotein nt sequence of IMMV-Ir (MN746770) based on the *de novo* assembly, were generated. There was no evidence for the presence of other *potyviruses* in the NGS contigs. To confirm the paired-end RNA sequencing results of the contig sequences, RT-PCR and Sanger sequencing were done with *potyvirus* degenerate primers for CI and HC-Pro regions (Ha *et al*., 2008) and virus-specific primers described by Li *et al*. (2016) (for ISMV) and Wylie *et al*. (2019) (for IMMV).

The genome of ISMV-Ir is composed of 10,454 nt, excluding the poly(A) tail, with the organization typical of *ISMV*. The 5′ and 3′ NCRs are 123 and 380 nt in length. The large ORF of 9,951 nt starts with an AUG codon (positions 124 to 126) and terminates with UAG codon (positions 10,072 to 10,074). This ORF encodes a polyprotein of 3,316 aa with a molecular mass of 377.26 kDa. The

small ORF PIPO exists in the middle of the P3 coding region. Its predicted start is the highly conserved G_1A_6 motif (nucleotides 3,640-3,646) beginning at nt position 3,642 and terminating at termination codon UAA at positions 3,882 to 3,884. The calculated mass of the PIPO is 9.99 kDa. The polyprotein of *ISMV*-Ir was processed into ten mature functional proteins in a manner like other *potyviruses*, and the nine putative protease cleavage sites were identified (Table 2). The previously identified and conserved functional motifs described in *potyviruses* (Urcuqui-Inchima *et al*., 2001; Adams *et al*., 2005; Moradi *et al*., 2017) were present in ISMV-Ir polyprotein. However, in the N-terminal of HC-Pro, the conserved motif KITC, associated with aphid transmission (Plisson *et al*., 2003) was altered to $K_{499}IGC_{502}$. Also, the conserved motif Asp-Ala-Gly (DAG), which is an important factor to regulate virus transmission by aphids (Atreya *et al*., 1995), was changed to $T_{3085}AG_{3087}$, in concordance with previous reports (Li *et al*., 2016; Van der Vlugt *et al*., 1994).

To date, the partial cds of polyprotein of only two IMMV isolates (lacking the most of P1 cistron) are available in the GenBank database: two Australian isolates (MH886513, JF320812).

The genome of the Iranian isolate was also incomplete and the attempts to obtain the 5′end and 3′-UTR of the IMMV-Ir genome failed. The obtained sequence was 9,159 nt in length, encodes a polyprotein of molecular mass 347.56 kDa (3,052 aa) (Table 3), which carries conserved motifs and domains, as were also present in other *potyviruses* (Adams *et al*., 2005). The typical size of PIPO for *IMMV*-Ir was confirmed as 79 aa.

The order and relative positions of the motifs in the IMMV-Ir -encoded polyprotein is similar to that observed in other potyviruses: the proteolytic domain, cysteine proteinase domain, RNA helicase domains, the nucleotide-binding motif, and RNA-dependent RNA polymerase motifs. IMMV lacks a DAG motif and contains an NVG motif instead.

Sequence Comparison, Genetic Diversity, and Phylogenetic Analysis

The primarily sequence analyses indicated that ISMV isolates were most closely related to OYDV group of *potyviruses* (49.50% average nt identity, 39.50% average aa identity). The genetic distance (d) value for the complete ORFs of four ISMV isolates (one from this study and three isolates from the GenBank database) was 0.074±0.002, indicating high genetic diversity (Table 4). The genomic sequence of ISMV-Ir shared identities of 92.73-94.10% and 96.77-97.41% with three other ISMV isolates available in the GenBank at the nucleotide (nt) and amino acid (aa) levels, respectively. ISMV-Ir displayed the highest nt (94.10%) and aa (97.41%) sequence identity with a Japanese isolate J (LC433737) and the lowest nt (92.73%) and aa (96.77%) sequence identity with a Chinese isolate BJ (KT692938). Isolate ISMV-Ir also shared 82.48-98.86% nt identity (86.75-100% aa identity) in the CP core coding region with 10 other *ISMV* isolates, for which only partial sequences are available (see Table 1). It displayed the greatest nt (98.86%) and aa (100%) sequence identity with TA4 (KT998900) and BJ (KT692938) from China; and the lowest nt (82.48%) and aa (86.75 and 87.18%) sequence identity with isolates K (AF034839, from Korea) and Crocus (X75939, from the Netherlands) in the cCP sequence. In addition, ISMV-Ir shared 100% aa identity (97.29% nt identity) with isolate J in the cCP region.

The genetic variation and polymorphism of the *ISMV* isolates were also computed for the complete ORF sequences and for each coding region separately using several genetic diversity parameters implemented in DnaSP6 (Table 1). Based on the complete ORF sequences, the Haplotype diversity (Hd) and genetic variation (π) for ISMV isolates were 1.000 and 0.07276, respectively, indicating a relatively high genetic diversity. The order of genetic variation of individual coding regions, from highest to lowest, was as

Table 2. Genomic structure of ISMV-Ir and putative cleavage sites.

Region	Start-end position in genome	Size in nt	Size in aa	Cleavage sites (C-terminus)
5′-UTR	1-123	123	-	-
P1	124-1452	1329	443	IEHY/S
HC-Pro	1453-2820	1368	456	YNVG/G
P3	2821-4230	1410	470	VSFQ/A
6K1	4231-4386	156	52	VRFE/A
CI	4387-6297	1911	637	VEFQ/G
6K2	6298-6456	159	53	VEFQ/G
NIa-VPg	6457-7020	564	188	PVIE/D
NIa-Pro	7021-7761	741	247	VTMQ/G
NIb	7762-9321	1560	520	VRFE/G
CP	9322-10074	753	250	-
3′-UTR	10075-10454	380	-	-

Table 3. Genomic structure of IMMV-Ir and putative cleavage sites.

Coding Region	Size in nt	Size in aa	Cleavage sites (C-terminus)
P1	69 (Incomplete)	23 (Incomplete)	FDLF/T
HC-Pro	2112	704	YRVG/G
P3	1044	348	VEHQ/A
6K1	159	53	VNHQ/S
CI	1929	643	VLHE/S
6K2	159	53	VNHQ/A
NIa-VPg	564	188	VDHE/A
NIa-Pro	726	242	MREQ/S
NIb	1560	520	VEHQ/S
CP	837	278	-

Table 4. Genetic diversity and genetic distance (d) values within the virus isolates.

	d	
Genomic region	ISMV isolates (n= 4)	IMMV isolates (n= 3)
Polyprotein	0.074±0.002	0.122±0.004
P1	0.075±0.005	- (incomplete)
HC-Pro	0.075±0.005	0.150±0.011
P3	0.069±0.006	0.134±0.011
PIPO	0.041±0.010	0.089±0.016
6K1	0.047±0.014	0.090±0.022
CI	0.066±0.005	0.103±0.006
6K2	0.067±0.016	0.125±0.028
NIa-VPg	0.046±0.007	0.141±0.015
NIa-Pro	0.051±0.006	0.135±0.013
NIb	0.138±0.014	0.104±0.013
CP	0.027±0.004	0.123±0.011
CP	(n= 11) 0.087+0.006	(n= 7) 0.109±0.010

follows: NIb, P1, HC-Pro, P3, 6K2, CI, NIa-Pro, 6K1, NIa-VPg, PIPO and CP (Table 1). The overall genetic distance among complete CP of 11 ISMV (including *ISMV*-Ir and 10 other isolates) was 0.087±0.006. The range of sequence identities among 11 ISMV CP gene was 81.15 to 100% (at nt level) and 84.52 to 100% (at aa level). The dN/dS (ω) ratio for the complete ORFs was less than 1 (dN/dS= 0.0519), indicating that the ISMV genome was under dominant purifying selection. To assess variation in selection pressure on each cistron, the dN/dS ratios were also calculated separately for each gene (Table 5). The results showed that the substitutions were not uniformly distributed along the ISMV genome. The strongest purifying selection was observed in the NIa-Pro protein, supported by the smallest ω value (0.0070), while the weakest purifying selection was in

Table 5. Genetic variation and selection pressure in polyprotein and individual genes of four iris severe mosaic virus isolates.

Genomic region	H	Hd	π	dN	dS	ω
Polyprotein	4	1.000	0.07276	0.01637	0.31506	0.0519
P1	4	1.000	0.07399	0.03374	0.22844	0.1476
HC-Pro	4	1.000	0.07373	0.01647	0.32580	0.0505
P3	4	1.000	0.06745	0.02184	0.27248	0.0801
PIPO	4	1.000	0.04085	0.02777	0.09331	0.2976
6K1	4	1.000	0.04665	0.00859	0.18138	0.0473
CI	4	1.000	0.06522	0.00865	0.29612	0.0292
6K2	4	1.000	0.06583	0.01208	0.32091	0.0376
NIa-VPg	4	1.000	0.04516	0.00227	0.22653	0.0100
NIa-Pro	4	1.000	0.05072	0.00175	0.24869	0.0070
NIb	4	1.000	0.13584	0.02637	0.91492	0.0288
CP	4	1.000	0.02706	0.00256	0.11904	0.0215

Table 6. Percent nucleotide and amino acid identities of each coding region of *Iris Severe Mosaic Virus* (ISMV) and *Iris Mild Mosaic Virus* (IMMV) isolates.

	ISMV isolates (n= 4)		IMMV isolates (n= 3)	
Genomic region	Percent nucleotide identity (Average)	Percent amino acid identity (Average)	Percent nucleotide identity (Average)	Percent amino acid identity (Average)
Polyprotein	92.08-94.05% (93.04)	96.04-97.40% (96.68)	76.17-90.07% (83.15)	83.59-91.09% (88.25)
P1	91.42-93.82% (92.82)	91.42-93.45% (92.73)	- (Incomplete)	- (Incomplete)
HC-Pro	91.73-94.07% (92.97)	94.95-97.36% (96.34)	50.18-79.88% (63.29)	52.41-84.86% (65.70)
P3	91.77-95.39% (93.55)	93.61-97.44% (95.24)	83.04-99.52% (88.63)	90.22-98.85% (93.29)
PIPO	94.16-98.33% (96.038)	91.25-97.50% (93.54)	87.76-99.15% (91.83)	75.94-98.73% (83.96)
6K1	92.94-99.35% (95.50)	96.15-100% (98.07)	88.05-99.37% (92.03)	94.33-100% (96.22)
CI	93.09-94.40% (93.75)	97.17-98.90% (98.00)	86.52-99.84% (90.97)	96.26-99.84% (97.50)
6K2	91.82-95.59% (93.70)	94.33-100% (97.16)	84.27-100% (89.51)	98.11-100% (98.74)
NIa-VPg	93.79-97.70% (95.62)	98.93-100% (99.46)	82.26-99.29% (88.17)	93.61-100% (95.74)
NIa-Pro	94.73-95.68% (95.09)	99.19-100% (99.59)	82.78-99.72% (88.51)	97.52-100% (98.34)
NIb	81.28-91.41% (87.68)	93.26-98.26% (95.95)	86.34-99.61% (90.82)	96.73-99.80% (97.81)
CP	96.94-98.53% (97.34)	99.60-100% (99.80)	75.63-99.47% (83.57)	78.41-98.39% (85.19)

the PIPO followed by the P1 protein, with the largest ω values (0.2976 and 0.1476, respectively). Pairwise sequence identities were analyzed at both the nucleotide and amino acid levels, either at genome scale or at individual gene scale (Table 6), which confirmed the genetic variation results. With respect to the separate genes, the highest nt identity among ISMV isolates was found in the CP gene (with the average nt identity 97.34%), followed by NIa-VPg and NIa-Pro; NIb had the lowest nt identity (average 87.68%), followed by P1 (92.82%) (Table 5).

At the amino acid level, the highest aa identity was found in CP (average 99.80%), followed by NIa-Pro (average 99.59%) and NIa-VPg (average 99.46%), while the lowest identity was found in P1 (average 92.73%) (Table 6). For IMMV isolates, the closest virus identified from genomic sequence analysis was *Chilli Veinal Mottle Virus* (ChVMV; 54.55% average nt identity, 52% average aa identity) and *Tobacco Vein Banding Mosaic Virus* (TVBMV, 54.65% average nt identity, 52.30% average aa identity).

The value of genetic distance (d) among partial polyprotein nt sequences of three *IMMV* isolates (one from this study and two isolates from GenBank) was very high (0.122±0.004), indicating that these isolates were genetically divergent (Table 4). The genomic sequence of IMMV-Ir isolate shared 82.36% nt identity (91.25% aa identity) with BC32 isolate (MH886513) from Australia. It also shared the lowest nt (75.55%) and aa (83.59%) sequence identity with WA-1 isolate (JF320812) from Australia. Isolate IMMV-Ir shared identities of 76.55–98.45% (nt) and 80.93–99.15% (aa) in the CP core coding region with six other IMMV isolates (see Table 1). The highest identity was found

with Australian isolate Bate2 (JN127338, with 98.31% nt and 99.15% aa identity) and with New Zealand isolate DC4b (DQ436919, with 98.45% nt and 98.73% aa identity); and the lowest nt (76.55%) and aa (80.93%) sequence identity were found with BC32 isolate (MH886513). The coat protein sequence of IMMV-Ar (KX870019) was not included in this study, as it was incomplete. However, IMMV-Ir and IMMV-Ar shared identities of 98.0 (at the nt level) and 97.8% (at the aa level) with each other in the partial NIb-CP sequences (702 bp). Analysis of nucleotide diversity and haplotype diversity revealed a high level of polymorphism for the IMMV polyprotein (π= 0.1171, Hd= 1.000). In comparison, across the coding regions (except for the P1 region that was incomplete), the greatest and the least values of genetic variation were found for HC-Pro (π= 0.14401) and 6K1 (π= 0.08581) coding regions, respectively (Table 7). The high level of genetic diversity (0.109±0.010) was also obtained within complete CP nucleotide sequences of seven IMMV isolates. The dominant selective pressure across *IMMV* genomic sequences was negative (ω< 1) (Table 7), which meant that the sequence variation along the IMMV isolates was controlled by purifying selection. The maximum and minimum dN/dS ratio was estimated for PIPO (ω= 0.9050) and NIa-Pro (ω= 0.0049) genes, respectively. Among putative gene products of IMMV (excluding the P1 cistron), HC-Pro was the most variable (65.70%), and the NIa-Pro (98.34%) and 6K2 (98.74%) were the most conserved proteins (Table 5).

Analysis of recombination in the full-length polyprotein nucleotide sequences of ISMV isolates led to the detection of three possible recombinants. ISMV-Ir was a putative recombinant with isolates ISMV-J and ISMV-BJ as its parental sequences and the putative recombination regions occurring at position 7335-8100 of polyprotein nt sequence, between the central part of NIa-Pro and the N-terminal of NIb. Furthermore, one possible recombination event was detected for ISMV-J isolate (located at position 7654 to 9298 of polyprotein nt sequence) and for *ISMV*-Tai'an isolate (located at position 7666 to 8631 of polyprotein nt sequence). The details about major and minor parents, programs detecting the event, and the *P*-values can be found in Table 8. However, due to limited sequences for analysis, these isolates seemed to be "tentative" recombinants. It is also worth noting that no recombination sites were found in the coding regions of IMMV isolates. A phylogenetic analysis was performed initially using the genomic sequences of ISMV and IMMV and 49 other members of the genus *Potyvirus* (data not shown). In addition, ML and NJ tree analyses of the amino acid sequences gave trees with the same topology (Figure 2). Likewise, the topologies of both ML and NJ trees of cCP coding sequences were essentially identical. Phylogenetic analyses, based on complete amino acid sequences or cCP coding sequences (705 nt) of ISMV and IMMV isolates and 49 additional *potyviruses*, confirmed the clustering of ISMV isolates with OYDV, SYSV, NDV, and VSV and the clustering and close relationship of IMMV with ChVMV and TVBMV (Figures 2 and 3). Among the 16 main *potyvirus* clades (represented by light and dark gray rectangles), 11 distinct lineages of *potyviruses* grouped by Gibbs and Ohshima (2010) were indicated (Figure 2). Members of the ChVMV group were divided into two subgroups, and IMMV-Ir along with WA-1 and BC32 isolates formed a separate subgroup (Figure 2).

Table 7. Genetic polymorphism in polyprotein and individual genes of three iris mild mosaic virus isolates.[a]

Genomic region	H	Hd	π	dN	dS	ω
Polyprotein[b]	3	1.000	0.11717	0.02527	0.68513	0.0368
P1	-	-	-	-	-	-
HC-Pro	3	1.000	0.14414	0.05701	0.66230	0.0860
P3	3	1.000	0.12849	0.03709	0.66252	0.0559
PIPO	3	1.000	0.08855	0.08651	0.09559	0.9050
6K1	3	1.000	0.08640	0.01598	0.53544	0.0298
CI	3	1.000	0.09936	0.01226	0.60519	0.0202
6K2	2	0.667	0.11764	0.00549	1.01543	0.0054
NIa-VPg	3	1.000	0.13416	0.02369	0.93066	0.0254
NIa-Pro	3	1.000	0.13013	0.01025	2.07931	0.0049
NIb	3	1.000	0.10098	0.01103	0.63292	0.0174
CP	3	1.000	0.11892	0.04251	0.52927	0.0803

[a] H= Number of Haplotypes; Hd= Haplotype diversity; π= nucleotide diversity; dS= Synonymous nucleotide diversity; dN= Non-synonymous nucleotide diversity; ω= dN/dS, average ratio between nonsynonymous and synonymous substitutions in sequence pairs.

[b] Incomplete ORF, due to the lack of P1 sequence data.

Table 8. Analysis of possible recombination in the full-length polyprotein nucleotide sequences of ISMV isolates by RDP4.

		ISMV-Ir	ISMV-J	ISMV-Tai'an
Major parent		J (LC433737, Japan)	BJ (KT692938, China)	J (LC433737, Japan)
Minor parent		BJ (KT692938, China)	Tai'an (MF385582, China)	BJ (KT692938, China)
P-values determined by different programs	RDP	1.943×10^{-17}	4.645×10^{-9}	1.060×10^{-43}
	GENECONV	2.808×10^{-17}	7.520×10^{-6}	3.960×10^{-43}
	BOOTSCAN	5.754×10^{-16}	8.277×10^{-9}	2.558×10^{-43}
	MAXCHI	3.895×10^{-13}	8.809×10^{-7}	6.174×10^{-20}
	CHIMERA	4.252×10^{-12}	2.401×10^{-8}	2.891×10^{-15}
	SISCAN	9.325×10^{-29}	2.100×10^{-19}	1.248×10^{-30}
	3SEQ	1.773×10^{-12}	-	4.440×10^{-16}
Beginning-ending breakpoint in polyprotein nt sequence		7335-8100	7654-9298	7666-8631

DISCUSSION

Deep sequencing revealed that symptomatic iris plants collected in eastern Iran in 2019, infected with at least two RNA viruses, belonged to two different *Potyvirus* species: ISMV and IMMV. It has been indicated that the DAG motif, which interacts with PTK of HC-Pro to regulate virus transmission by aphids, is not universally conserved (Nigam *et al*., 2019). As identified in this study, DAG in the coat protein of ISMV and IMMV changed to TAG and NVG, respectively. Consistent with these results, several *potyviruses* that are transmitted by aphids do not contain a DAG motif (Johansen *et al*., 1996; Wylie *et al*., 2002; Nigam *et al*., 2019). According to previous studies, certain substitutions in the DAG motif or proximal residues cause loss or appreciable reduction in the aphid transmissibility of tobacco vein mottling virus (Atreya *et al*., 1990, 1991). However, not all *potyviruses* with this motif are aphid transmissible and certain modifications of the DAG are tolerated by the virus (Johansen *et al*., 1996; Ateka *et al*., 2017) and these deviations of DAG do not affect CP-HC-Pro binding (Moradi *et al*, 2017). All the evidence together suggests a correlation between variation in CP, aphid-transmission, and virus-vector specificity (Dombrovsky *et al*., 2005; Nigam *et al*., 2019). Another possible explanation for this modification is the potential transmission of these viruses by infected iris rhizomes, which indicates the less importance of aphid transmission.

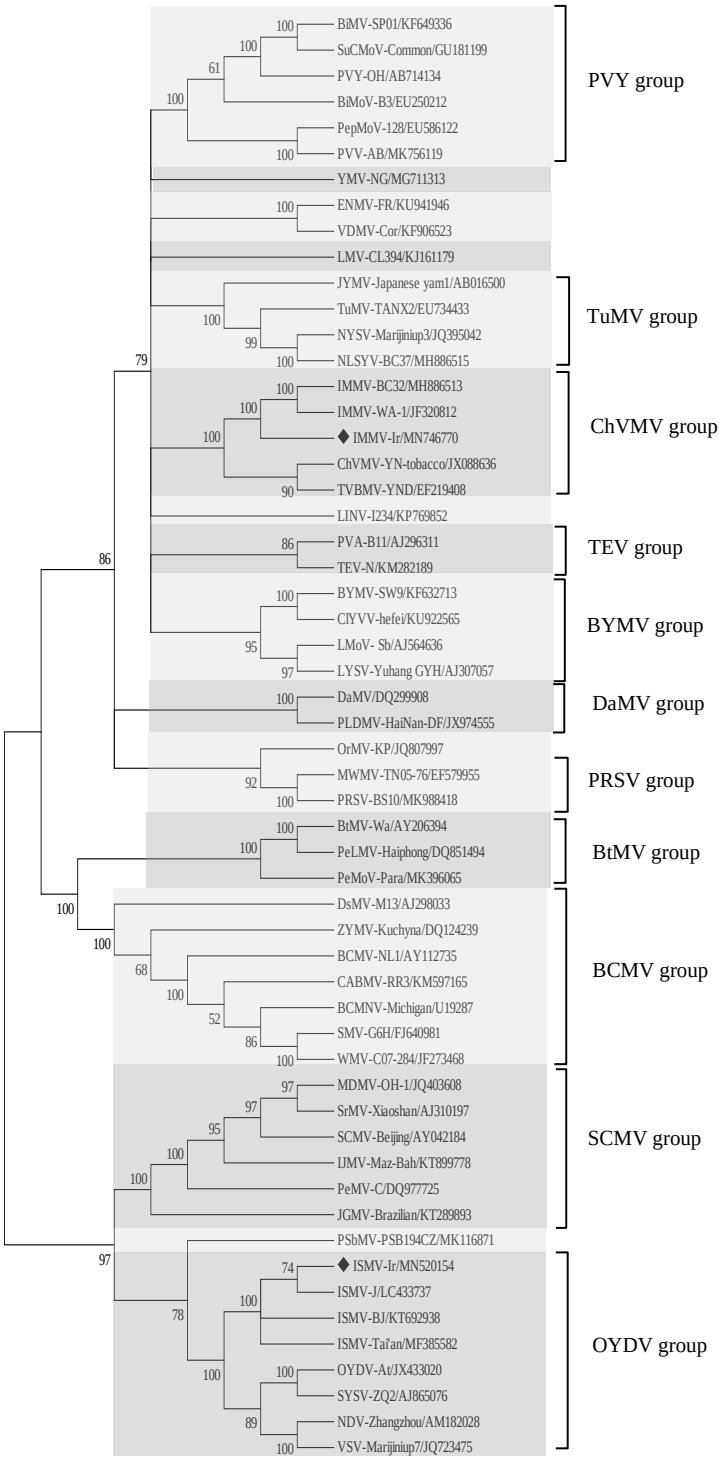

Figure 2. Maximum likelihood tree based on the polyprotein amino acid sequences of ISMV (n= 4) and IMMV (n= 3) isolates compared with the corresponding region of 49 other *potyviruses*. Bootstrap support (n= 1000) above 50% is indicated for each node. The scale bar represents the number of substitutions per amino acid. The clades are represented by light and dark gray rectangles and 11 distinct groups of *potyviruses* according to Gibbs and Ohshima (2010) clustering are indicated on the tree. Iranian ISMV and IMMV isolates were marked. Information about each virus is described in detail in Table 1.

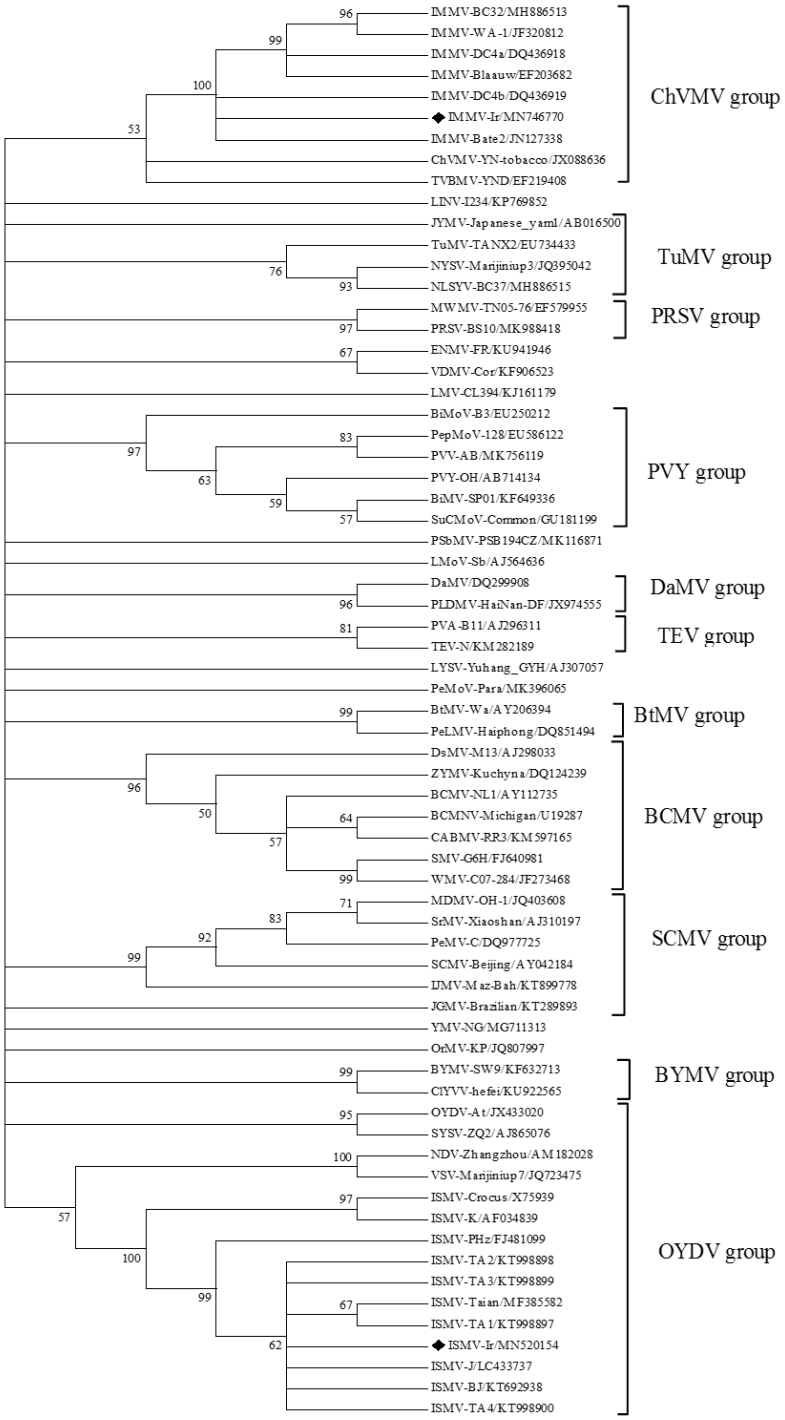

Figure 3. Maximum likelihood tree based on core coat protein coding sequences of ISMV and IMMV isolates and those of 49 other *potyviruses*. Bootstrap supports (n= 1000) above 50% are indicated for each node.

Gibbs and Ohshima (2010) classified 59 *potyviruses* into eleven groups and seven singletons based on the analyses of polyproteins encoded by their genomes. Recently, Moury and Desbiez (2020) established a new clade outside the 11 groups (totally 12 main clades), based on the analyses of amino acid sequences of 59 *potyviruses*. In this study, the ML and NJ trees (based on polyprotein sequences or cCP coding sequences) clearly placed *ISMV* and IMMV in a distinct well-supported branch within the OYDV and ChVMV groups, respectively (Figures 2 and 3). Despite the use of a different set of *potyvirus* species, the topology of the trees agreed with those obtained by Gibbs and Ohshima (2010) and by Moury and Desbiez (2020) (Figure 2). These results were in agreement with pairwise sequence comparisons and BlastN and BlastX results. ISMV-Ir had 92.73 to 94.10% nt sequence identity and 96.77 to 97.41% aa sequence identity with three other ISMV isolates available in the GenBank, the highest with Japanese isolate J (LC433737). IMMV-Ir had 75.55-82.36% nt sequence identity and 83.59-91.25% aa sequence identity with two other IMMV isolates available in the GenBank, the highest with Australian isolate BC32 (MH886513). The results of genetic distance and polymorphism analyses were consistent with the pairwise sequence comparisons (Table 6). The divergence value of ISMV polyprotein nt sequences was 0.074, which was reduced by purifying selection pressure, as has been reported for other *potyviruses* (Nigam *et al*., 2019; Moradi and Mehrvar, 2019). NIb exhibited the most genetic divergence (d= 0.138±0.014), compared to P1 (d= 0.075±0.005), HC-Pro (d= 0.075±0.005) and other coding regions of ISMV (Table 4). However, the low dN/dS ratio in NIb (0.0288) is consistent with a strong negative selection against protein change. The high negative selection pressure on the NIb, NIa (Vpg and Pro), CI, and CP of the ISMV can be attributed to the functional importance of these proteins during infection and replication (Revers and Garcia, 2015; Cheng and Wang, 2017; Ivanov *et al*., 2014; Dombrovsky *et al*., 2005; Deng *et al*., 2015; Martinez *et al*., 2016). The result also showed less negative selection pressure (ω= 0.1476) acting on the P1 relative to most of the cistrons. On the other hand, between putative gene products of ISMV, P1 was the most variable protein (92.73%, see Table 6), followed by P3 (95.24%), NIb (95.95%), and HC-Pro (96.34%). P1 diversification can be attributed to the host adaptation in a wide range of host species, host-dependent pathogenicity, vector transmissibility, and viral accumulation in different hosts (Johansen *et al*., 1996; Tan *et al*., 2005; Valli *et al*., 2007; Moury and Simon, 2011).

A lower divergence value in the CP cistron of ISMV (Table 4) indicates greater sequence conservation in the capsid protein (99.80% average aa identity), as depicted in Table 6. A highly significant genetic divergence (0.177±0.026) was found for the polyprotein of IMMV isolates. The presence of a high level of variation in the HC-Pro (63.29% nt, 65.70% aa) and CP (83.57% nt, 85.19% aa) proteins may be an indication of hypervariability of these regions, as observed for other *potyviruses* (Nigam *et al*., 2019). This could be explained by the participation of both CP and HC-Pro in aphid transmission and virus-vector-host specificity (Dombrovsky *et al*., 2005). Additionally, the presence of indels in these regions may be another reason for the greater sequence divergence. For all coding genes, the dN/dS ratio was less than one, which is consistent with purifying selection to maintain the functional integrity of the viral genome. The NIa-Pro gene had the lowest mean dN/dS ratio, which meant that this region was under strong purifying selection against aa changes. This is consistent with the high degree of conservation observed in NIa-Pro protein (98.34% average aa identity), which strongly indicates the functional significance of this protein. Although our analysis revealed three possible recombination events for ISMV isolates and not for IMMV, these findings may not be accurate and reliable due to the small number of available sequences. Hence,

additional genomic sequences from other countries are needed to obtain an accurate prediction of recombination and full assessment of the genetic variability of these viruses. Co-infection of ISMV with IMMV and other viruses such as cucumber mosaic virus has been reported (Van der Vlugt, 1994; Asjes, 1979). Therefore, the generation of new viral strains by recombination, an important evolutionary force that has contributed to the divergence of several positive-sense RNA viruses (particularly potyviruses) (Chare and Holmes 2006), can be a threat for the iris plants.

Our studies showed that ISMV-Ir and IMMV-Ir were not mechanically transmitted from a symptomatic iris plant to a number of plant species. This is in line with other studies (Derks and Hollinger, 1986, Van der Vlugt *et al*., 1994). Iris plants are propagated mainly from vegetative propagation organs such as bulbs or creeping rhizomes, which are important sources for the virus (Van der Vlugt *et al*., 1994). As ISMV and IMMV are vertically transmitted through vegetative propagation, their entry in Iran through contaminated propagation material cannot be ruled out. The economic impact of these viruses can be important, hence, to prevent the severe loss, appropriate management strategies need to be devised.

In conclusion, this is the first comprehensive report of the nearly complete genomic sequence and organization of two distinct *potyvirus* isolates coinfecting the iris plant in Iran, with high confidence. Our results provide valuable information to facilitate the development of strategies for the control of iris *potyviruses*.

REFERENCES

1. Adams, M. J., Antoniw, J. F., Beaudoin, F. 2005. Overview and Analysis of the Polyprotein Cleavage Sites in the Family *Potyviridae*. *Mol. Plant. Pathol.,* **6:** 471–487.
2. Asjes, C. J. 1979. Viruses and Virus Diseases in Dutch Bulbous Irises (*Iris hollandica*) in the Netherlands. *Neth. J. Plant Pathol.,* **85:** 269-279.
3. Ateka, E., Alicai, T., Ndunguru, J., Tairo, F., Sseruwagi, P., Kiarie, S., Makori, T., Kehoe, M. A. and Boykin, L. M. 2017. Unusual Occurrence of a DAG Motif in the *Ipomovirus Cassava Brown Streak Virus* and Implications for Its Vector Transmission. *PLoS One,* **12:** e0187883.
4. Atreya, C.D., Raccah, B. and Pirone, T. P. 1990. A Point Mutation in the Coat Protein Abolishes Aphid Transmissibility of a *Potyvirus*. *Virology*, **178:** 161–165.
5. Atreya, P. L., Atreya, C. D., Pirone, T. P. 1991. Amino Acid Substitutions in the Coat Protein Result in Loss of Insect Transmissibility of a Plant Virus. *Proc. Natl. Acad. Sci. USA*, **88:** 7887±91.
6. Atreya, P.L., Lopez-Moya, J. J., Meihua Chu, Atreya, C. D. and Pirone, T. P. 1995. Mutational Analysis of the Coat Protein N-Terminal Amino Acids Involved in *Potyvirus* Transmission by Sphids. *J. Gen. Virol.,* **76:** 265-270.
7. Chare, E. R. and Holmes, E. C. 2006. A Phylogenetic Survey of Recombination Frequency in Plant RNA Viruses. *Arch. Virol.,* **151:** 933-946.
8. Chen, J., Shi, Y. H., Li, M. Y., Adams, M. J. and Chen, J. P. 2008. A New *Potyvirus* from Butterfly Flower (*Iris japonica* Thunb.) in Zhejiang, China. *Arch. Virol.,* **153:** 567–569.
9. Cheng, X. and Wang, A. 2017. The *Potyvirus* Silencing Suppressor Protein vpg Mediates Degradation of SGS3 *via* Ubiquitination and Autophagy Pathways. *J. Virol.,* **91:** e01478-16.
10. Deng, P., Wu, Z. and Wang, A. 2015. The Multifunctional Protein CI of *Potyviruses* Plays Interlinked and Distinct Roles in Viral Genome Replication and Intercellular Movement. *Virol. J.,* **12:** 141.
11. Derks, A. F. L. M. and Hollinger, T. C. 1986. Similarities of and Differences between *Potyviruses* from Bulbous and Rhizomatous Irises. *ISHS Acta Hortic.,* **177:** 555–561.
12. Dombrovsky, A., Huet, H., Chejanovsky, N. and Raccah, B. 2005. Aphid Transmission of a *Potyvirus* Depends on Suitability of the Helper Component and the N Terminus of the Coat Protein. *Arch. Virol.,* **150:** 287–298.
13. Gibbs, A. and Ohshima, K. 2010. *Potyviruses* and the Digital Revolution. *Annu. Rev. Phytopathol.,* **48:** 205-223.

14. Ha, C., Coombs, S., Revill, P. A., Harding, R. M., Vu, M. and Dale, J. L. 2008. Design and Application of Two Novel Degenerate Primer Pairs for the Detection and Complete Genomic Characterization of *Potyviruses*. *Arch. Virol.*, **153:** 25-36.
15. Ivanov, K. I., Eskelin, K., Lohmus, A. and Makinen, K. 2014. Molecular and Cellular Mechanisms Underlying *Potyvirus* Infection. *J. Gen. Virol.,* **95:** 1415-1429.
16. Johansen, I. E., Keller, K. E., Dougherty, W. G. and Hampton, R. O. 1996. Biological and Molecular Properties of a Pathotype P-1 and a Pathotype P-4 Isolate of *Pea Seed-Borne Mosaic Virus. J. Gen. Virol.,* **77:** 1329–1333.
17. Kulshrestha, S., Hallan, V., Raikhy, G., Kumar, A., Ram, R., Garg, I. D. and Zaidi, A. A. 2004. Molecular Characterization of an *Iris Severe Mosaic Virus* Isolate from India. *Acta Virol.,* **48:**65–67.
18. Kulshrestha, S., Hallan, V., Raikhy, G., Ram, R., Zaidi, A. A. and Garg, I. D. 2006a. Incidence of *Bean Yellow Mosaic Virus* in Iris. *Acta Hortic.,* **722:** 235–240.
19. Kulshrestha, S., Hallan, V., Raikhy, G., Ram, R., Garg, I. D., Haq, Q. M. R. and Zaidi, A. A. 2006b. Occurrence of *Iris Mild Mosaic Potyvirus* in Cultivated Iris in India. *Indian J. Biotechnol.,* **5:** 94–98.
20. 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.
21. Li, Y., Deng, C., Shang, Q., Zhao, X., Liu, X. and Zhou, Q. 2016. The First Complete Genome Sequence of *Iris Severe Mosaic Virus*. *Arch. Virol.,* **161:** 1069-1072.
22. 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.
23. Martinez, F., Rodrigo, G., Aragones, V., Ruiz, M., Lodewijk, I., Fernandez, U., Elena, S. F. and Daros, J. A. 2016. Interaction Network of *Tobacco Etch Potyvirus* NIa Protein with the Host Proteome during Infection. *BMC Genom.,* **17:** 87.
24. Moradi, Z., Mehrvar, M., Nazifi, E. and Zakiaghl, M. 2017. *Iranian Johnsongrass Mosaic Virus*: The Complete Genome Sequence, Molecular and Biological Characterization, and Comparison of Coat Protein Gene Sequences. *Virus Genes,* **53:** 77-88.
25. 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.
26. Moury, B. and Simon, V. 2011. dN/dS-Based Methods Detect Positive Selection Linked to Trade-offs between Different Fitness Traits in the Coat Protein of *Potato Virus Y. Mol. Biol. Evol.,* **28:** 2707–2717.
27. Moury, B. and Desbiez, C. 2020. Host Range Evolution of Potyviruses: A Global Phylogenetic Analysis. *Viruses,* **12:** 111.
28. Nateqi, M., Habibi, M. K., Dizadji, A. and Parizad, S. 2015. Detection and Molecular Characterization of the *Iris Severe Mosaic Virus-Ir* Isolate from Iran. *J. Plant Prot. Res.,* **55:** 235-240.
29. Nateqi, M., Dizadji, A., Koohi Habibi, M. and Movi, S. 2017. First Report of *Iris Mild Mosaic Virus* from *Iris xiphium* in Iran. *J. Plant Pathol.,* **99:** 287–304.
30. Nigam, D., LaTourrette, K., Souza, P. F. N. and Garcia-Ruiz, H. 2019. Genome-Wide Variation in *Potyviruses*. *Front. Plant Sci.,* **10:** 1439.
31. Nishikawa, M., Tokuda, R., Yoshida, T., Nijo, T., Maruyama, N., Katsu, K., Maejima, K., Yamaji, Y. and Namba, S. 2019. Complete Genome Sequence of *Iris Severe Mosaic Virus* Isolated in Japan. *Microbiol. Resour. Announc.,* **8:** e00093-19.
32. Park, W. M., Lee, S. S., Park, S. H. and Ryu, K. H. 2000. Sequence Analysis of the Coat Protein Gene of a Korean Isolate of Iris Severe Mosaic Potyvirus from Iris Plant. *Plant Pathol J.,* **16:** 36–42.
33. Plisson, C., Drucker, M., Blanc, S., German-Retana, S., Le Gall, O., Thomas, D. and Bron, P. 2003. Structural Characterization of HC-Pro, a Plant Virus Multifunctional Protein. *J. Biol. Chem.,* **278:**23753–61.
34. Revers, F. and Garcia, J. A. 2015. Molecular Biology of Potyviruses. *Adv. Virus Res.,* **92:** 101-199.
35. Rodamilans, B., Valli, A., Mingot, A., San León, D., Baulcombe, D., López-Moya, J. J. and García, J. A. 2015. RNA Polymerase Slippage as a Mechanism for the Production of Frameshift Gene Products in Plant Viruses of the *Potyviridae* Family. *J. Virol.,* **89:** 6965–6967.

36. Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E. and Sanchez-Gracia, A. 2017. DnaSP6: DNA Sequence Polymorphism Analysis of Large Datasets. *Mol. Biol. Evol.,* **34:** 3299-3302.
37. Seguin, J., Rajeswaran, R., Malpica-López, N., Martin, R. R., Kasschau, K., Dolja, V. V., Otten, P, Farinelli, L. and Pooggin, M. M. 2014. *De Novo* Reconstruction of Consensus Master Genomes of Plant RNA and DNA Viruses from siRNAs. *PLoS One,* **9:** e88513.
38. Stevens, P. F. 2001 *Onwards. Angiosperm Phylogeny.* Website. http://www.mobot.org/MOBOT/research/APweb/.
39. Tan, Z., Gibbs, A. J., Tomitaka, Y., Sanchez, F., Ponz, F. and Ohshima, K. 2005. Mutations in *Turnip Mosaic Virus* Genomes that Have Adapted to *Raphanus sativus.* J. Gen. Virol., **86:**501–510.
40. Urcuqui-Inchima, S., Haenni, A. and Bernardi, F. 2001. Potyvirus Proteins: A Wealth of Functions. *Virus Res.,* **74:** 157-175.
41. Valli, A., López-Moya, J. J. and García, J. A. 2007. Recombination and Gene Duplication in the Evolutionary Diversification of P1 Proteins in the Family *Potyviridae*. *J. Gen. Virol.,* **88:** 1016–1028.
42. Van der Vlugt, C. I. M., Derks, A. F. L. M., Boonekamp, P. M. and Goldbach, R. W. 1993. Improved Detection of *Iris Severe Mosaic Virus* in Secondarily Infected Iris Bulbs. *Ann. Appl. Biol.,* **122:** 279–288.
43. Van der Vlugt, C. I. M., Langeveld, S. A. and Goldbach, R. W. 1994. Molecular Cloning and Sequence Analysis of the 3'-Terminal Region of *Iris Severe Mosaic Virus* RNA. *Arch. Virol.,* **136:** 397–406.
44. Van der Vlugt, C. I. M. 1994. Distribution and Multiplication of *Iris Severe Mosaic Potyvirus* in Bulbous Iris in Relation to Metabolic Activity: Implications for ISMV Detection. Ph.D. Thesis, Bulb Research Centre Lisse, Dutch Flower bulb Industry, Netherland, 130 PP.
45. Wei, T., Pearson, M. N. and Cohen, D. 2006. First Report of *Ornithogalum Mosaic Virus* and *Ornithogalum Virus 2* in New Zealand. *Plant Pathol.,* **55:** 820.
46. Wei, T., Pearson, M. N. and Cohen, D. 2007. First Report of *Narcissus Latent Virus* in New Zealand. *Plant Pathol.,* **56:**720.
47. Wylie, S. J., Kueh, J., Welsh, B., Smith, L. J., Jones, M. G. K. and Jones, R. A. C. 2002. A Non-Aphid-Transmissible Isolate of *Bean Yellow Mosaic Potyvirus* Has an Altered NAG Motif in Its Coat Protein. *Arch. Virol.,* **147:** 1813–1820.
48. Wylie, S. J., Luo, H., Li, H. and Jones, M. G. K. 2012. Multiple Polyadenylated RNA Viruses Detected in Pooled Cultivated and Wild Plant Samples. *Arch. Virol.,* **157:** 271–284.
49. Wylie, S. J., Adams, M., Chalam, C., Kreuze, J., López-Moya, J. J., Ohshima, K., Praveen, S., Rabenstein, F., Stenger, D., Wang, A. and 2017. Zerbini, F. M. ICTV Report Consortium. ICTV Virus Taxonomy Profile Potyviridae. *J. Gen. Virol.,* **98:** 352-354.
50. Wylie, S. J., Tran, T. T., Nguyen, D. Q., Koh, S. -H., Chakraborty, A., Xu, W., Jones, M. G. K. and Li, H. 2019. A Virome from Ornamental Flowers in an Australian Rural Town. *Arch. Virol.,* **164:** 2255–2263.
51. Yan, S. J., Qin, Z. D., Jin, L. L. and Chen, J. S. 2010. A New Isolate of *Iris Severe Mosaic Virus* Causing Yellow Mosaic in *Iris ensata* Thunb. *J. Nanosci. Nanotechnol.,* **10:** 726–730.

مشخصات ژنومی و تجزیه و تحلیل فیلوژنتیکی دو *پوتی ویروس* آلوده کننده زنبق در ایران

ع. ناصری، ز. مرادی، م. مهرور، و م. زکی‌عقل

چکیده

چندین ویروس گیاه زنبق را آلوده می‌کنند و به دلیل خسارتهای جدی اقتصادی، از محدودیت‌های عمده در تولید تجاری این گیاه هستند. در این بررسی، اولین توالی ژنومی دو پوتی‌ویروس به نام‌های ویروس موزائیک شدید زنبق (ISMV) و ویروس موزاییک خفیف زنبق (IMMV) از *Iris versicolor* در ایران که به طور طبیعی آلوده شده بودند با استفاده از توالی‌یابی عمیق RNA و به دنبال آن با RT-PCR تعیین شد. هر دو ویروس (ISMV-Ir, IMMV-Ir) دارای سازمان ژنتیکی تیپیک موجود در پوتی‌ویروس‌ها بوده، و از یک قاب خواندنی باز (ORF) که کدکننده یک پلی‌پروتئین حاوی نه مکان برش اتوکاتالیستی بوده و نیز یک پروتئین فرضی کوچکتر به نام P3N-PIPO تشکیل شده‌اند. تجزیه و تحلیل فیلوژنتیک و مقایسه توالی‌ها بیانگر روابط نزدیک بین ISMV و اعضای گروه OYDV از Potyvirusها بود. ISMV-Ir بیش از ۹۲٪ تشابه نوکلئوتیدی (بیش از ۹۶٪ تشابه آمینواسیدی) با سه جدایه دیگر ISMV از دیگر کشورها نشان داد. ISMV-Ir بیشترین تشابه (۹۴/۱۰٪ در سطح نوکلئوتیدی و ۹۷/۴۱٪ در سطح آمینواسیدی) را با جدایه ژاپنی J و کمترین تشابه (۹۲/۷۳٪ در سطح نوکلئوتیدی و ۹۶/۷۷٪ در سطح آمینواسیدی) را با جدایه چینی BJ داشت. IMMV-Ir متعلق به گروه ChVMV از پوتی‌ویروس‌ها بوده، و دارای ۸۲/۳۶٪ تشابه نوکلئوتیدی (۹۱/۲۵٪ تشابه آمینواسیدی) با جدایه BC32، و ۷۵/۵۵٪ تشابه نوکلئوتیدی (۸۳/۵۹٪ تشابه آمینواسیدی) با جدایه WA-1 از استرالیا بود. فاصله ژنتیکی بین توالی‌های ژنومی کدکننده پلی‌پروتئین IMMV یا توالی‌های اختصاصی هر ژن نشان‌دهنده واگرایی ژنتیکی بالای این جدایه‌ها است. آنالیزهای ما نشان داد که انتخاب طبیعی در تکامل جدایه‌های متعلق به این دو پوتی‌ویروس دخیل بوده است. اطلاعات مربوط به توالی‌های ژنومی ارائه شده در این مطالعه درک ما را از عملکرد و بیماری زایی ویروس بهبود می بخشد که در نهایت منجر به کنترل بهتر بیماری ناشی از آنها می‌گردد.