

A New Distinct Clade for Iranian *Tomato spotted wilt virus* Isolates Based on the Polymerase, Nucleocapsid, and Non-structural Genes

Mahsa Abadkhah, Davoud Koolivand *, and Omid Eini

Department of Plant Protection, Faculty of Agriculture, University of Zanjan, Zanjan, Iran

(Received on April 16, 2018; Revised on May 30, 2018; Accepted on July 24, 2018)

Tomato spotted wilt virus (TSWV; Genus *Orthospovirus*: Family *Tospoviridae*) is one of the most destructive viruses affecting a wide range of horticultural crops on a worldwide basis. In 2015 and 2016, 171 leaf and fruit samples from tomato (*Solanum lycopersicum*) plants with viral symptoms were collected from the fields in various regions of Iran. ELISA test revealed that the samples were infected by TSWV. The results of RT-PCR showed that the expected DNA fragments of about 819 bp in length were amplified using a pair of universal primer corresponding to the RNA polymerase gene and DNA fragments of ca 777 bp and 724 bp in length were amplified using specific primers that have been designed based on the nucleocapsid (N) and non-structural (NSs) genes, respectively. The amplified fragments were cloned into pTG19-T and sequenced. Sequence comparisons with those available in the GenBank showed that the sequences belong to TSWV. The high nucleotide identity and similarities of new sequences based on the L, N, and NSs genes showed that minor evolutionary differences exist amongst the isolates. The phylogenetic tree grouped all isolates six clades based on N and NSs genes. Phylogenetic analysis showed that the Iranian isolates were composed a new distinct clade based on a part of polymerase, N and NSs genes. To our knowledge, this is the first detailed study on molecular

characterization and genetic diversity of TSWV isolates from tomato in Iran that could be known as new clade of TSWV isolates.

Keywords : Neighbor-Joining, *Orthospovirus*, phylogenetic analysis, *Tomato spotted wilt virus*

Handling Associate Editor : Ju, Ho-Jong

Tomato (*Solanum lycopersicum*) is one of the most important vegetable crops in the world (Kimura and Sinha, 2008) and is cultivated in different regions of Iran. Tomato production is impacted due to many viral diseases which substantially reduce yield and quality (Golnaraghi et al., 2004).

Orthospovirus is the only genus of arthropod-vectoring viruses in the family *Tospoviridae* whose members infect different host plant such as tomato, pepper, groundnut, and numerous ornamental plants (10th report ICTV: <https://talk.ictvonline.org/taxonomy/>). *Orthospovirus*'s particles are quasi-spherical, 80-120 nm in diameter and enveloped by a lipid membrane (Chen et al., 2016). The type member of this genus is *Tomato spotted wilt virus* (TSWV) which has been studied extensively because it has economic impact and broad host range in several important crops especially tomato (Sherwood et al., 2003).

TSWV causes serious damage to a wide range of economically important crops worldwide including fruits, vegetables, ornamental, and weed plants as well (Parrella et al., 2003). TSWV isolates are transmitted by several species of thrips (*Thysanoptera*, Thripidae) in a persistent and propagative manner, especially the western flower thrips *Frankliniella occidentalis* (Jones, 2005; Parrella et al., 2003; Whitfield et al., 2005). Various symptoms including chlorosis, necrosis and ring spots on leaves, fruit, and stems that lead to wilting and death of infected plants are more common in TSWV infection (Tripathi et al., 2015).

*Corresponding author.

Phone) +98 24 3305 2465, FAX) +98 24 3228 3202

E-mail) Koolivand@znu.ac.ir

ORCID

Davoud Koolivand

<http://orcid.org/0000-0002-5863-2591>

© 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.

The genome of TSWV consists of three single stranded RNA segments denoted as L (8.9 kb), M (4.8 kb), and S (2.9 kb) RNAs which together encode five proteins (Kaye et al., 2011). The L RNA is of negative polarity and encodes a RNA polymerase on the viral complementary strand, which plays an important role in viral replication. In contrast, the other two genomic RNAs use an ambisense coding strategy. The M RNA encodes a nonstructural (NSm) protein in the viral sense and a protein to serve as the precursor for the G_N and G_C glycoproteins in the viral complementary sense. The NSm protein is involved in cell-to-cell movement of non-enveloped ribonucleocapsid structures. The S RNA encodes in the viral sense, a nonstructural (NSs) protein that functioned as RNA silencing suppressor and in the viral complementary sense, a nucleocapsid protein (NP) that encapsidates viral RNAs (King et al., 2011; Margaria et al., 2015; Takeda et al., 2002).

Tomato spotted wilt virus was first described in 1919 on tomato in Australia (Brittlebank, 1919; Samuel et al., 1930) and since has been reported from many other countries (Lian et al., 2013). Recently, TSWV was detected in Middle-Eastern and Far-Eastern countries in Asia including Japan, South Korea, and China (Zhang et al., 2016).

Tomato spotted wilt virus was reported from Iran based on symptoms and electron microscopy in 1996 (Bananej et al., 1996). There are some reports of Orthotospoviruses from different parts of Iran in tomato, potato, soybean, and ornamental plants based on serological methods and host plants reactions (Ghotbi and Shahraeen, 2012; Golnaraghi et al., 2007; Hajiabadi et al., 2012; Pourrahim et al., 2001).

Limited information is available on the occurrence and sequence characteristics of TSWV in tomato fields from Iran. It is clear that survey for TSWV infection, sequence analysis of the virus and differentiation of viral isolates provide useful information for genetic characterization of the isolates based polymerase gene, NSs, and N genes. Hence, the aim of this study was to investigate the occurrence and genetic diversity of TSWV isolates in tomato fields in Iran, and determine the evolutionary model and genetic relationships of Iranian isolates compared to known TSWV isolates.

Materials and Methods

Virus Source. During spring through fall in 2015 and 2016 around 20 tomato fields were visited to collect the suspected samples with TSWV symptoms. Samples (N=171) were collected based on symptoms such as spots, chlorosis, malformations, necrosis on leaves, and spots on fruits from tomato fields in different regions from West and North-

Table 1. Reaction of the indicator plants to TSWV infection after mechanical inoculation

Indicator plants	Reactions
Cucurbitaceae	
<i>Cucumis sativus</i>	CL
<i>Cucurbita pepo</i>	CL
Fabaceae	
<i>Vigna unguiculata</i>	NL, CS, LD
Solanaceae	
<i>Solanum lycopersicum</i> var <i>Early Urbana</i>	CL, NL, GR, VC
<i>Solanum lycopersicum</i> var <i>Superchief</i>	CL, NL, GR, VC
<i>Nicotiana benthamiana</i>	CS, NL, LD
<i>N. tabacum</i> cv. <i>Samsun</i>	CS, NL, LD

CL = chlorotic lesions, CS = chlorotic spots, GR = growth reaction, LD = leaf deformation, NL = necrotic lesion, VC = veinal chlorosis

western of Iran. A part of samples was stored at -80°C for RNA extraction, and fresh samples were inoculated on different plants for under greenhouse conditions.

Experimental Host range. Host range studies and symptomatology were carried out under greenhouse conditions. The host species that were inoculated included *Cucumis sativum* (Cucumber), *Cucurbita pepo* (Squash), *Vigna unguiculata* (Cowpea), *Solanum lycopersicum* var *Early Urbana* and, *S. lycopersicum* var *Superchief* (Tomato), *Nicotiana benthamiana* and, *N. tabacum* cv. *samsun* (Tobacco) (Table 1).

Serological Assay for detection of TSWV. Initially, double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) (Clark and Adams, 1977) was used using polyclonal antibodies (Agdia, USA) against TSWV (diluted 1:1000). An absorbance value at (405 nm) was measured with using an ELISA microplate reader (BioTek ELX 8000, USA). The threshold was set as twice the mean absorbance value of the healthy plant sap.

RNA extraction and RT-PCR. Total RNA was extracted from 100-200 mg of leaves or fruit tissue samples using RNX-plus solution (SinaClon, Iran) according to manufacturer's instruction and finally suspended in 25 μl sterile distilled water. Optical density (OD) measurements of each extraction were taken to quantify the RNA concentration (ng/ μl) and purity (at 260/280 and 260/230 wavelength ratios) using the NanoDrop 2000 (Thermo Scientific, USA). Then, the reverse transcription reaction of 1 μg total RNA was done using Hyperscript master mix (Genall, South Korea) with random hexamer primer in a total volume of 10 μl , according to manufacturer's instruction.

Table 2. Primers used in this research

Primers	Sequence	Region
TSWVN-F	5' ATCGGATCCATGTCTAAGGTTAAGCTCAC 3'	N
TSWVN-R	5' ATCCTCGAGTTAAGCAAGTTCTGTGAGTTTTGC 3'	N
NSS2-F	5' GCTTCAGTCTGGGGATCAACT 3'	NSs
NSS2-R	5' TTGGAACCTTAGCCAGAGGC 3'	NSs
gl3637-F	5' CCTTTAACAGT(A/T/G)GAAACAT 3'	RdRp
gl4435c-R	5' CAT(A/T/G)GC(A/G)CAAGA(A/G)TG(A/G) TA(A/G)ACAGA 3'	RdRp

To test the presence of TSWV, a pair of universal primers (gl3637/gl4435c) corresponded to a conserved region in *Orthospovirus* polymerase L gene (Chu et al., 2001) was used for the amplification of an 819 bp fragment. Additionally, two pair primers were designed for RT-PCR detection of TSWV (Table 2), for the N and NSs genes of TSWV to amplify DNA fragments of about 777 bp and 724 bp, respectively. Primer pairs were designed by aligning several sequences of TSWV taken from GenBank (www.ncbi.nlm.nih.gov) using ClustalW alignment in MEGA version 7 (Table 2).

PCR products were amplified under the following conditions: 2 min at 94°C; 35 cycles of 30 s at 94°C, 45 s at 50°C (for universal primers) or 45 s at 52°C (For N and NSs genes primers), and 1 min at 72°C. A final polymerization step at 72°C for 10 min was also applied in three cases.

Cloning. Amplified fragments resulting from the reverse transcription polymerase chain reaction (RT-PCR) were ligated into the pTG19-T vector (Vivantis, Malaysia) using 1 Unit of T4 DNA ligase (Vivantis, Malaysia) according to the manufacturer's protocol and incubated at 4°C overnight. Ligation mixes were transformed into prepared chemically competent *Escherichia coli* strain DH5α cells by heat shock method (Chung et al., 1989). The transformed cells were selected on LB plates containing ampicillin (100 mg/ml), IPTG (100 µl of 0.1 M per plate), and X-Gal (20 µl of 50 mg/ml). Plasmids were extracted by Accuprep nano-pluse plasmid mini extraction kit (Bioneer, South Korea) according to manufacturer's instruction. The recombinant plasmids were verified by PCR colony and each colony carrying the cloned cDNA was subjected to single colony isolation. Also, the purified plasmids were digested by *Bam*HI enzyme (Thermo Fisher Scientific, USA) to release the DNA inserted and loaded on 1% agarose gel. Next, the independent clones of each viral isolate were subjected to nucleotide sequencing with M13F/R primers. The sequencing was done by Bioneer Inc. (Seoul, South Korea).

Phylogenetic analysis. The obtained sequences (Nucleotide and amino acid) were analyzed and compared with those available in the GenBank using BLAST program. The sequences were aligned with TSWV isolates around the world (Table 3) by MEGA7 and algorithm of ClustalW and the distance matrix based on Jukes and Cantor's model was used to estimate nucleotide divergence. The phylogenetic relationships were determined with Neighbor-Joining (NJ) (Saitou and Nei, 1987). The robustness of the inferred evolutionary relationships was assessed by 1000 bootstrap replicates for phylogenetic analysis based on the part of L, N, and NSs genes. All branches with < 50% bootstrap were collapsed. Nucleotide identity and similarity were determined using the MEGA7 software. In addition, Sequence Demarcation Tool version 1.2 (SDT v1.2) (Muhire et al., 2014) was used to generate pairwise nucleotide sequence identity matrix.

Genetic diversity analysis. DnaSP version 6.10.01 (Rozas et al., 2017) was used to provide the confidence intervals of the number of haplotypes (H), haplotype diversity (Hd), number of polymorphic (segregation) sites (S), total number of mutations η (Eta), average number of nucleotide differences (k), average pairwise nucleotide diversity (π), total number of synonymous sites (SS), total number of non-synonymous sites (NS), and the ratio of non-synonymous nucleotide diversity to synonymous nucleotide diversity (ω) known as $\omega=dN/dS$. There are three types of selection pressure, including negative (purifying), neutral, and positive (diversifying). The gene is under positive, neutral, and negative selection when ω ratio is > 1 , $= 1$ and < 1 , respectively (Rozas et al., 2017). To calculate individual codon positions under natural selection was used of single likelihood ancestor counting algorithm (SLAC) which defined in the free and online Datamonkey webserver (<http://www.datamonkey.org>) within HyPhy software package (Pond and Muse, 2005).

Neutrality, genetic differentiation and gene flow statistical tests. To investigate the neutral selection hypothesis

Table 3. Origins, hosts, and accession numbers of *Tomato spotted wilt virus* isolates/strains analyzed in this research

Segment	Origin (Number of isolates)	Isolates/strains	Host	Accession number
S segment	South Korea (5)	LS3, TSWV-10, NJ-JN, TSWV-16, TSWV-8	<i>Leonurus sibiricus</i> , <i>Stellaria aquatic</i> , <i>Solanum lycopersicum</i> , <i>S. lycopersicum</i> , <i>Lactuca indica</i>	KM076653, KC261964, HM581936, KC261970, KC261961
	USA (4)	NC-3, CA-5, CA-2, BasC	<i>Dahlia</i> , <i>Chrysanthemum</i> , <i>Buttercup</i> , <i>Ocimum basilicum</i>	AY744478, AY744472, AY744469, KU179515
	Iran (4)	TOS101 (N, NSs), TOS102 (N, NSs), TSWV-TP1 (N, NSs), TSWV-RT (N, NSs)	(<i>S. lycopersicum</i> , <i>S. lycopersicum</i>), (<i>S. lycopersicum</i> , <i>S. lycopersicum</i>), (<i>S. lycopersicum</i> , <i>S. lycopersicum</i>), (<i>S. lycopersicum</i> , <i>S. lycopersicum</i>)	(KT899947, MG987616), (KT899948, MG987618), (KY923205, MG987617), (KY923204, KY923210)
	France (4)	LYE89 (N, NSs), LYEPO (N, NSs), STM3B (N, NSs), SOO6 (N, NSs)	(<i>S. lycopersicum</i> , <i>S. lycopersicum</i>), (<i>S. lycopersicum</i> , <i>S. lycopersicum</i>), (<i>Stellaria media</i> , <i>S. media</i>), (<i>Sonchus oleraceus</i> , <i>S. oleraceus</i>)	(FR693062, FR692838), (FR693063, FR692839), (FR693257, FR693033), (FR693256, FR693032)
	Italy (6)	p170, p202/3WT, p202, CHM6 (N, NSs), T-1003 (N, NSs), Miz-3 (N, NSs)	<i>Pepper</i> , <i>Pepper</i> , <i>Pepper</i> , (<i>Chrysanthemum murifolium</i> , <i>C. murifolium</i>), (<i>L. esculentum</i> , <i>Lycopersicon esculentum</i>), (<i>S. melongena</i> , <i>Solanum melongena</i>)	DQ431237, HQ830187, DQ398945, (FR693048, FR692824), (GU369725, GU369752), (GU369720, GU369746)
	Spain (5)	Pujol1TL3, SPAIN-2, VE427, LL-N.05, PVR	<i>S. lycopersicum</i> , <i>S. lycopersicum</i> , <i>Pepper</i> , <i>S. lycopersicum</i> , <i>Pepper</i>	KP008131, AY744480, DQ376185, KP008129, KP008134
	China (7)	YNgp, KM-T, TSWV-LE, CG-1, TSWV-Celery, YN5573, YN5574	<i>Green pepper</i> , <i>S. lycopersicum</i> , <i>lettuce</i> , <i>Lactuca sativa</i> , <i>Celery</i> , <i>Pea</i> , <i>Codonopsis pilosula</i>	KM657116, HQ402595, KU976396, JN664252, KU356854, KY495609, KY495610
L segment	South Korea (7)	TSWV-10, NJ-JN, TSWV-16, TSWV-8, Pepper1 CY-CN, Pepper2 CY-CN	<i>L. sibiricus</i> , <i>S. aquatic</i> , <i>S. lycopersicum</i> , <i>S. lycopersicum</i> , <i>L. indica</i> , <i>Pepper</i> , <i>Pepper</i>	KC261962, HM581934, KC261968, KC261959, HM581937, HM581940
	China (6)	YNgp, TSWV-LE, CG-1, YNrp, TSWV-YN, YNta	<i>G. pepper</i> , <i>lettuce</i> , <i>L. sativa</i> , <i>Red pepper</i> , <i>S. lycopersicum</i> , <i>Tobacco</i>	KM657122, KU976394, JN664254, KM657120, JF960237, KM657121
	Spain (3)	Pujol1TL3, LL-N.05, PVR	<i>S. lycopersicum</i> , <i>S. lycopersicum</i> , <i>Pepper</i>	KP008130, KP008128, KP008132
	Italy (2)	p202/3WT, p105	<i>Pepper</i> , <i>Pepper</i>	KJ575619, KJ575620
	Iran (4)	T-3/55(TOS101), TOS-301 (TOS102), TOS-401(TSWV-TP1), TOS-501(TSWV-RT)	<i>S. lycopersicum</i> , <i>S. lycopersicum</i> , <i>S. lycopersicum</i> , <i>S. lycopersicum</i> ,	KY923206, KY923207, KY923208, KY923209

operating by Tajima's D (Tajima, 1989), Fu and Li's D* & F* (Fu and Li, 1993) statistical tests and to calculate statistical tests of population differentiation including K_S^* , K_{ST}^* , Z^* , S_{nn} and F_{ST} (Hudson, 2000; Hudson et al., 1992) between phylogroups and geographical populations were done using DnaSP v.6.10.01 for both N and NSs genes. Tajima's D test static is based on the differences between the number of segregating sites and the average number of

nucleotide differences (Tajima, 1989). Fu and Li's D* test is based on the differences between the number of singletons and the total number of mutation. Fu and Li's F* test statistic is based on the differences between the number of singletons and the average number of nucleotide differences between pairs of sequences (Fu and Li, 1993). If there is no genetic differentiation (under the null hypothesis), K_{ST}^* is expected to be near zero, but if it is supported by a small

P value (< 0.05) the null hypothesis is rejected (Tsompana et al., 2005). The Z^* static is an algorithmic variant of the Z static. Also, if it is too small and supported by significant P value (< 0.05) the null hypothesis is rejected (Hudson et al., 1992). The nearest neighbor of sequences is measured by the Snn test static, whose P value is ranged between 1 (when population is distinctly differentiated) to one-half in the case of panmixia (Hudson, 2000). Finally, the coefficient of F_{ST} (genetic differentiation) was used to estimate inter-population diversity and the absolute value of F_{ST} ranges between zero to one for indicating undifferentiated to fully differentiated populations (Hudson et al., 1992; Tsompana et al., 2005). Normally, $F_{ST} > 0.25$ suggests a large gene flow (genetic differentiation) within the populations (Gao et al., 2016).

Results

Field observation and host range. During the tomato-growing season, virus symptoms including spots on the fruits and leaves, stem necrosis, chlorosis, leaf deformation and yellow spot discoloration on ripe fruits and subsequent wilting and complete collapse of plants were observed in the inspected fields (Fig. 1) that based on virus descriptions, they seemed to be infected by TSWV. Generally, necrotic spots were observed at young stages of tomato plants on leaves. At the same time that leaf symptoms were developed in the fields, fruit symptoms such as yellow spot were

observed in some fields (Fig. 1). It should be noted that some symptoms such as mild and sever mosaic, leaf deformation and shoe string were observed in the fields with common symptoms of TSWV in some samples. A wide range of the observed symptoms in the fields may occur due to mixed infection with other viruses such as *Cucumber mosaic virus* (CMV) and *Tobacco mosaic virus* (TMV). Despite the fact that we expected to see the large population of vectors (different species of thrips) in the fields, the vectors observed in some fields as a small population.

The result of mechanical inoculation of plants with sap from field samples showed a range of symptoms at 10-20 days post inoculation (dpi) in different hosts (Table 1). *Cucumis sativus* and *Cucurbita pepo*, showed chlorotic lesions after 10 to 15 dpi; whereas systemic symptoms such as leaf deformation, mottling, and necrotic lesions were seen on *Vigna unguiculata* after 20 dpi (Fig. 2). Necrotic local lesions on *Solanum lycopersicum* (var *Early Urbana* and *Superchief*), and chlorotic spots on *Nicotiana benthamiana* and *N. tabacum* cv. *samsun* were visible 15 to 20 dpi (Fig. 2).

Serological analysis. TSWV was detected by DAS-ELISA in infected leaves and fruits samples from several tomato fields. This infectivity assay provided additional evidence as to the identity of the infecting virus being TSWV.

RT-PCR. The results of RT-PCR using universal primers

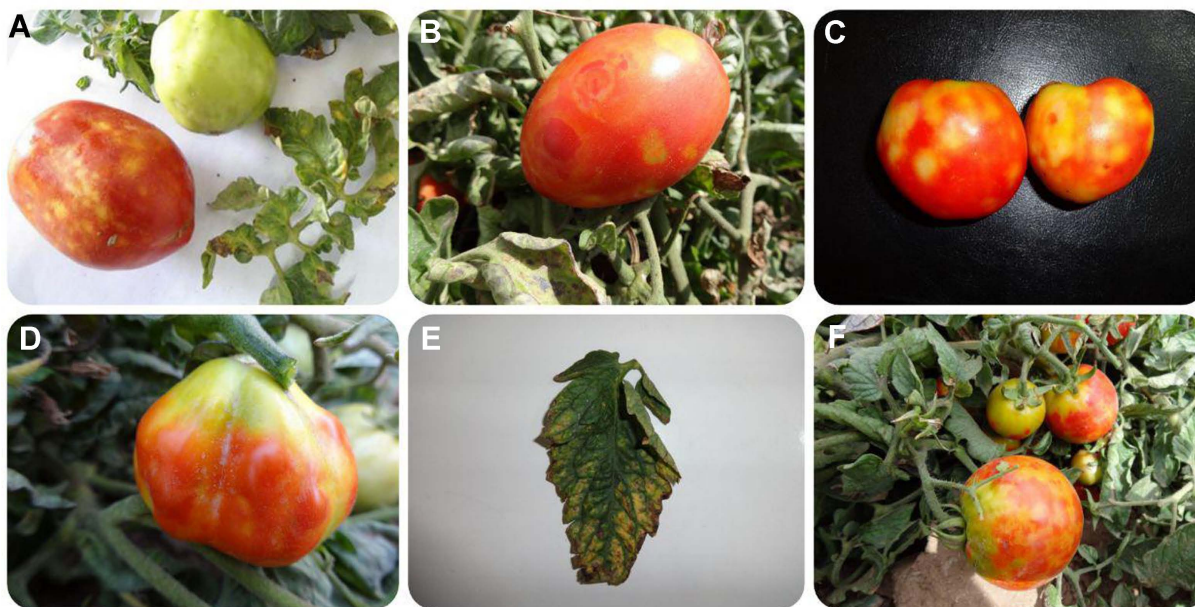


Fig. 1. Virus symptoms observed on tomato plants. (A) and (E) chlorotic and necrosis spots on leaf and fruit; (B) ring spots on fruit; (C) chlorotic blotches on fruits; (D) deformity and chlorotic spots on fruit; (F) chlorotic spots on fruits.



Fig. 2. Symptoms production in plants inoculated with TSWV. (A) and (B) *Solanum lycopersicum*, (C) *Nicotiana benthamiana*, (D) *N. tabacum* cv *samsun*, (E) *Vigna unguiculata*, (F) *Cucurbita pepo*.

showed that the expected DNA fragment of about 819 bp in length corresponding to a part of the polymerase gene (L segment) in 46 out of 171 samples. In addition, a DNA fragment about 777 bp encompassing the complete N gene was amplified in 14 out of 46 samples using specific primers. No amplification was obtained from healthy samples that were used as a negative control. Also, by using specific primers corresponding to a part of the NSs gene a DNA fragment of about 724 bp was amplified in 10 out of 14 samples. No DNA fragment was amplified from the healthy control. Optimal amplification conditions included annealing temperature of 50°C, 52°C, and 52°C for amplification of polymerase, N, and NSs genes, respectively. No

DNA fragment was amplified from the samples that were negative in DAS-ELISA.

Phylogenetic analysis. Blast analysis of our new nucleotide sequences (Table 3) with those in the GenBank, based on the high levels of sequence similarities, revealed that the new sequences belonged to TSWV. The identities of a part of L polymerase gene were 94-99% between the Iranian and other reported isolates in GenBank whereas the identities among the new TSWV isolates from Iran were 98% to 99% based on nucleotide sequences. Also, alignment of deduced amino acid sequences showed that the identities between the Iranian isolates and other isolates were around

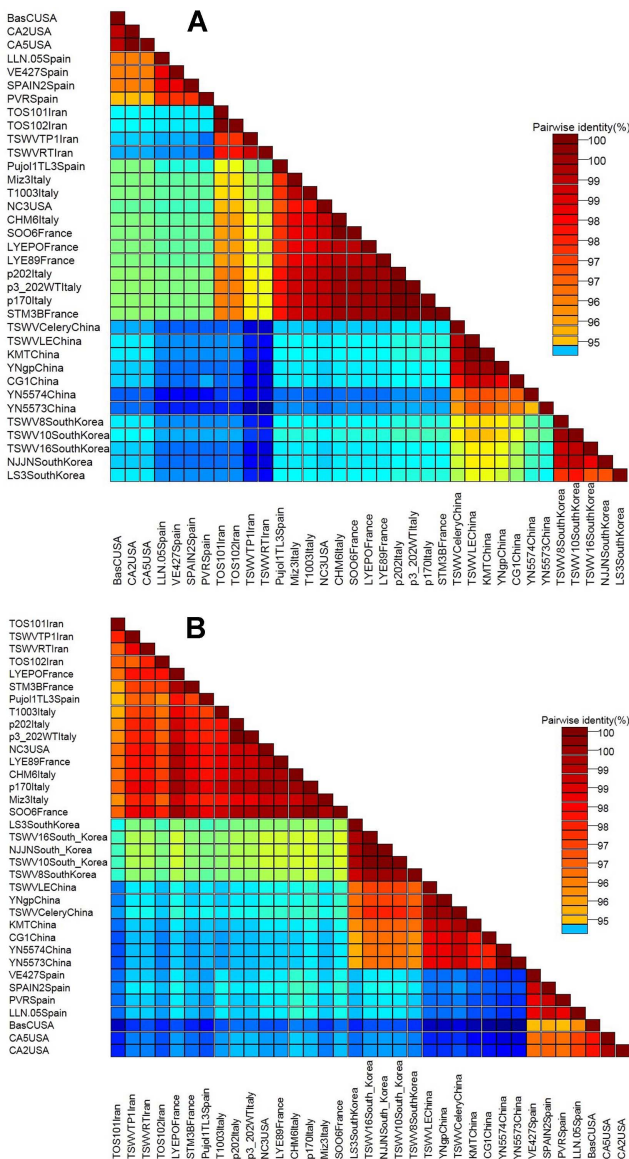


Fig. 3. Pairwise nucleotide sequence identity matrix of TSWV isolates from Iran and representative isolates from the GenBank, generated using SDT software for nucleocapsid gene (A) and non-structural gene (B).

98-99%.

For N gene, the identity of nucleotide and amino acid sequences of the complete N gene for the new Iranian isolates showed 99% and 98% identity, respectively. The Iranian isolates were most closely to six Italian isolates (Table 3), one isolate from USA (AY744478), and two isolates from France (FR692838, FR692839) at 98.3-98.8% nucleotide sequence identity when the new Iranian isolates compared to isolates from elsewhere in the world based on the complete N gene. Italian isolate showed the highest similarity (98.8%) with Iranian isolates (Fig. 3A).

The nucleotide sequence of partial NSs gene of the new Iranian isolates showed 94-99.3% identities to isolates around the world with most closely related to three isolates from France (FR692838, FR692839, FR693033), and two isolates from Italy (DQ431237, DQ398945) at 98.7-99.4% nucleotide sequence identity. The sequences that obtained from a part of NSs gene in this study showed high sequence identity (98.8-100%) among themselves (Fig. 3B). Pairwise nucleotide sequence identity matrix showed the identity result as matrix (Fig. 3B).

The phylogenetic tree based on a part of L segment consisted of five main clusters (I, II, III, IV and V) that were divided again into several branches (Fig. 4). The cluster I contained the Euro-Asian isolates from South Korea and Spain. The new Iranian sequences were grouped a distinct cluster in Cluster II. All isolates in cluster III originated from China. Other main clusters (IV and V) contained Asian isolates (South Korea), and European isolates (Spain and France) as two distinct clusters (Fig. 4).

The result of phylogenetic analysis based on the nucleocapsid gene for 31 isolates around the world and four new

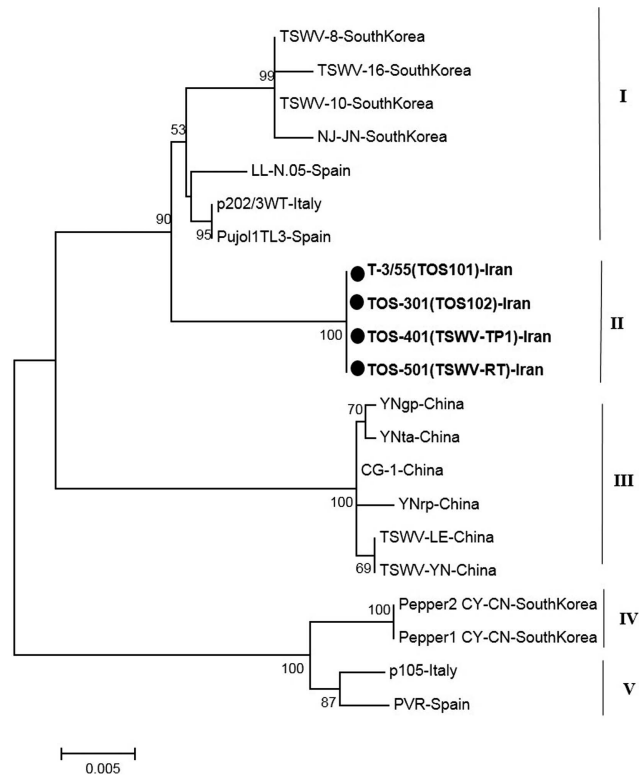


Fig. 4. Phylogenetic analysis based on the nucleotide sequences of a part of polymerase gene generated using Neighbor-Joining method by Mega 7. Bootstrap values on the branches represent the percentages out of 1000 bootstrap replicates program and Iranian isolates detected in this study have been marked by circle.

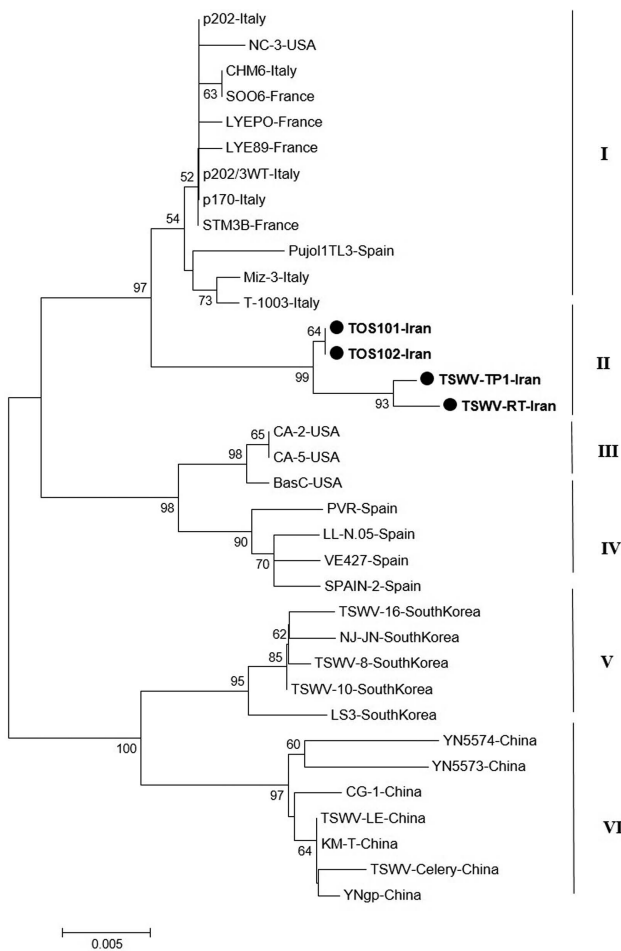


Fig. 5. Phylogenetic analysis based on the nucleotide sequences of the nucleocapsid gene generated using Neighbor-Joining method by Mega 7. Bootstrap values on the branches represent the percentages out of 1000 bootstrap replicates program and Iranian isolates detected in this study have been marked by circle.

isolates from Iran with Neighbor-Joining (NJ) method based on the nucleotide sequences produced six independent clusters namely I, II, III, IV, V and VI (Fig. 5). Cluster I include isolates from Italy, Spain, France, and USA; cluster II include new isolates from Iran as a distinct cluster. The main isolates in cluster III were isolates from USA and four Spanish isolates were grouped in cluster IV. South Korean and Chinese isolates were formed two independent clusters V and VI based on the N gene. As noted before, the Iranian isolates formed a single clade distinct from other isolates based on the N gene and were more closely related to isolates from Italy and Spain.

The phylogenetic tree based on nucleotide sequences of the NSs gene of TSWV isolates including four new isolates from Iran produced six main clusters consisting of cluster I with isolates from Europe (Spain, France, and Italy) and

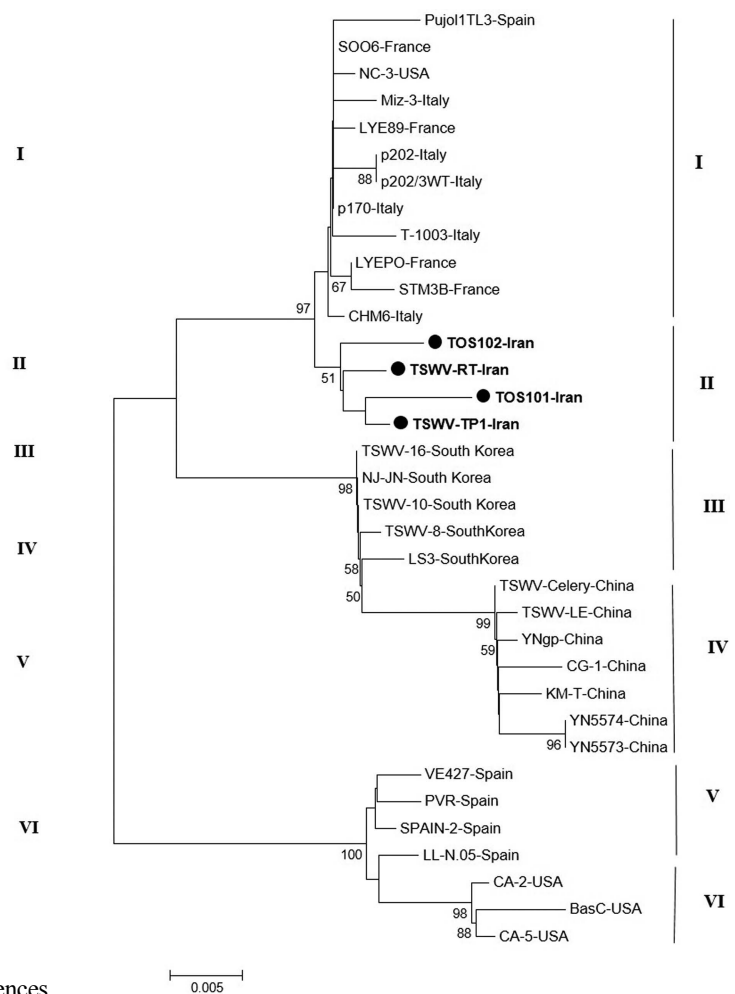


Fig. 6. Phylogenetic analysis based on the nucleotide sequences of the non-structural gene generated using Neighbor-Joining method by Mega 7. Bootstrap values on the branches represent the percentages out of 1000 bootstrap replicates program and Iranian isolates detected in this study have been marked by circle.

one isolate from USA. The Iranian isolates were grouped in independent cluster II that are near to isolates from in cluster I. Cluster III and IV are consisted of all isolates from East-Asian that South Korean isolates were grouped in cluster III and cluster IV is include Chinese isolates. Cluster V include isolates from Spain and USA isolates were grouped cluster VI based on the NSs gene sequences (Fig. 6).

The phylogenetic tree based on the sequences of N and NSs genes of isolates was grouped all isolates in six main groups include cluster I, II, III, IV, V, and VI (Fig. 7) as two other trees based on the N and NSs genes independently. The new Iranian isolates were grouped into a distinct cluster II (Fig. 7). Cluster III include all isolates from USA. The isolates from Spain were grouped into cluster IV (Fig.

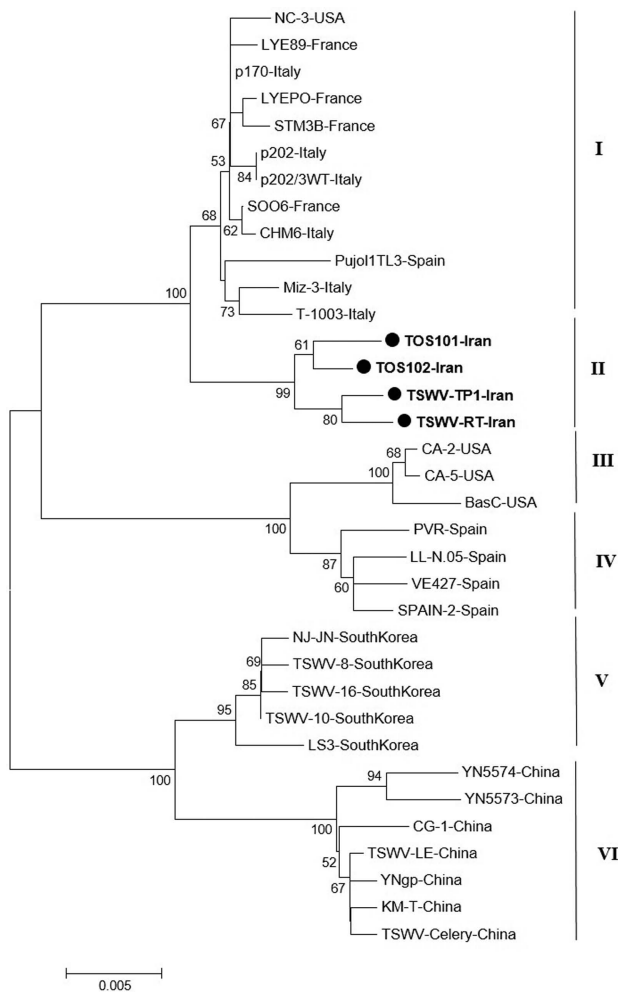


Fig. 7. Phylogenetic analysis based on the nucleotide sequences of nucleocapsid and non-structural genes generated using Neighbor-Joining method by Mega 7. Bootstrap values on the branches represent the percentages out of 1000 bootstrap replicates program and Iranian isolates detected in this study have been marked by circle.

7). Like other tree East-Asian isolates were formed two independent clusters that cluster V originated from South Korea and cluster VI were formed Chinese isolates.

Population genetic and polymorphism of TSWV. To obtain genetic variety of the TSWV populations based on the N and NSs sequences, several genetic diversity parameters were calculated (Tables 4, 5). The largest average numbers of differences, K (11 and 16 nucleotides) and also the greatest overall nucleotide diversity, π (0.015 and 0.023) between TSWV isolates were calculated for the Spanish and USA populations for N and NSs genes, respectively (Tables 4, 5). Also, the largest number of segregating sites S (28 and 32), and mutations within the segregating sites

η (28 and 32) were found in Spanish and USA TSWV populations for N and NSs genes, respectively (Tables 4, 5). Also, the smallest π (0.002 and 0.002) and k (2 and 1) were estimated for French and South Korean populations for N and NSs genes, respectively. Furthermore, the ratio of dN/dS (ω) was < 1 for all populations. The maximum ω ratios were obtained for TSWV from Iran (0.484) and China (0.329) for N and NSs genes, respectively, whereas the minimum ω values were calculated from the France (0 and 0) populations in both N and NSs genes. The results showed that both genes of TSWV populations are under negative selections. The highest (0.484) and lowest (0) ω values were calculated for clades II and III, respectively based on N gene. Also, clade V and clade I were introduced as highest (0.329) and lowest (0.101) ω values in NSs gene, respectively. The level of polymorphism analysis by “Sliding window” options window (50) and step size (25) were high at nucleotides 26-75, 51-100, 201-250, and 526-575 belong to the N gene with the π values of 0.039, 0.038, 0.040, and 0.051, and about NSs gene, were high at nucleotides 26-75, 101-150, and 454-503 with the π values of 0.041, 0.042, and 0.049 respectively (Figs. 8A, 8B, Supplementary). Based on the N gene there were no codons under positive selection by SLAC method in HyPhy software package which implemented in Datamonkey web-server with the *p-value* threshold ($p \leq 0.1$) (Fig. 9, Supplementary). This result showed that the nucleocapsid gene in TSWV is under strong negative evolutionary constraints.

Neutrality tests of TSWV. To explore the molecular variation patterns from segregated sites of TSWV populations based on the N and NSs sequences, several test statistics such as Tajima’s D, Fu and Li’s D* and F* (without an out group) were done (Tables 6, 7). Results showed that significantly negative values were obtained only for clade I in the NSs gene and non-significantly positive values were obtained for other clades in both genes (Tables 6, 7). Among geographic populations, non-significantly positive values were found only for Iranian population in both genes in all three test statistics and significantly negative value was calculated only for Spanish population in Fu and Li’s D* statistical test in NSs gene (Tables 6, 7).

Gene flow and genetic differentiation of TSWV populations. Gene flow and genetic differentiation analysis of TSWV populations showed that six phylogroups in both genes (N and NSs) with significant K_S^* , K_{ST}^* , and Z^* are completely distinct. Also, the related Snn values were significantly high (mostly 1.000 and/or near 1.000) (Tables 8, 9). The F_{ST} value between phylogroups was > 0.67 and

Table 4. Summary of genetic diversity and polymorphism analyses of TSWV N gene from different populations

Phylogroup	N	H	Hd	S	η	K	π	SS	NS	dS	dN	ω
All	35	28	0.983	92	93	18.071	0.02335	165.21	608.79	0.08854	0.00565	0.0638
Clade I	12	8	0.894	13	13	2.682	0.00346	165.11	608.89	0.01221	0.00109	0.0892
Clade II	4	3	0.833	7	7	4.167	0.00538	165.67	608.33	0.00905	0.00438	0.4839
Clade III	3	2	0.667	2	2	1.333	0.00172	165.17	608.83	0.00807	0.00000	0
Clade V	4	4	1.000	10	10	5.000	0.00646	164.92	609.08	0.02729	0.00082	0.0300
Clade VI	5	5	1.000	10	10	4.000	0.00517	164.73	609.27	0.01699	0.00197	0.1159
Clade VII	7	6	0.952	18	18	5.619	0.00726	165.67	608.33	0.01895	0.00407	0.2147
Geographic regions												
China	7	6	0.952	18	18	5.619	0.00726	165.67	608.33	0.01895	0.00407	0.2147
Iran	4	3	0.833	7	7	4.167	0.00538	165.67	608.33	0.00905	0.00438	0.4839
South Korea	5	5	1.000	10	10	4.000	0.00517	164.73	609.27	0.01699	0.00197	0.1159
USA	4	3	0.833	20	20	10.167	0.01314	165.17	608.83	0.05247	0.00246	0.0468
Spain	5	5	1.000	28	28	11.400	0.01473	164.83	609.17	0.05826	0.00295	0.0506
Italy	6	4	0.800	5	5	2.067	0.00267	165.06	608.94	0.01050	0.00055	0.0523
France	4	4	1.000	3	3	1.500	0.00194	165.33	608.67	0.00907	0.00000	0

N, number of isolates; H, number of haplotypes/isolates; 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; SS, total number of synonymous sites analyzed; NS, total number of non-synonymous sites analyzed; dS, synonymous nucleotide diversity; dN, non-synonymous nucleotide diversity. Numbers in bold are maximum and minimum values between populations.

Table 5. Summary of genetic diversity and polymorphism analyses of TSWV NSs gene from different populations

Phylogroup	N	H	Hd	S	η	K	π	SS	NS	dS	dN	ω
All	35	30	0.990	83	86	18.000	0.02651	152.86	522.14	0.08391	0.00991	0.1181
Clade I	12	10	0.970	17	17	3.242	0.00475	153.86	527.14	0.01566	0.00158	0.1008
Clade II	4	4	1.000	11	11	6.667	0.00978	152.92	528.08	0.02398	0.00568	0.2368
Clade III	5	3	0.700	3	3	1.200	0.00176	154.63	526.37	0.00518	0.00076	0.1467
Clade V	7	6	0.952	10	10	3.429	0.00505	153.33	521.67	0.01056	0.00347	0.3285
Clade VI	4	4	1.000	7	7	3.500	0.00513	153.67	527.33	0.01627	0.00190	0.1167
Clade VII	3	3	1.000	6	6	4.000	0.00587	153.00	528.00	0.01307	0.00379	0.2899
Geographic region												
China	7	6	0.952	10	10	3.429	0.00505	153.33	521.67	0.01056	0.00347	0.3285
Iran	4	4	1.000	11	11	6.667	0.00978	152.92	528.08	0.02398	0.00568	0.2368
South Korea	5	3	0.700	3	3	1.200	0.00176	154.63	526.37	0.00518	0.00076	0.1467
USA	4	4	1.000	32	32	16.000	0.02346	153.21	527.79	0.07172	0.00948	0.1321
Spain	5	5	1.000	32	32	12.800	0.01877	153.70	527.30	0.05984	0.00683	0.1141
Italy	6	5	0.933	8	8	3.067	0.00450	153.89	527.11	0.01560	0.00126	0.0807
France	4	4	1.000	4	4	2.167	0.00318	153.83	527.17	0.01408	0.00000	0

> 0.30 for N and NSs genes, respectively. The highest F_{ST} values (0.874 and 0.900) were obtained for South Korea versus France populations for N and NSs genes, respectively, and the lowest F_{ST} values (0.027 and 0.052) were found when comparing the France and Italy populations (for N gene) and the USA and Spain populations (for NSs gene), respectively (Tables 8, 9). Among TSWV populations based on the NSs gene, non-significant Z^* values were observed between Iran population with USA population, and France population with USA and Italian populations

(Table 9). Also, non-significant Z^* value was obtained between France population with Italy population based on the N gene (Table 8). Furthermore, among Iranian populations, non-significant S_{nn} value was found between Iran and USA populations and the highest F_{ST} value (0.856) was calculated from Iran population with South Korea population in comparison with other Iranian populations in the N gene, and based on the NSs gene, non-significant S_{nn} values were obtained between Iran population with USA and French populations and the highest F_{ST} value (0.808)

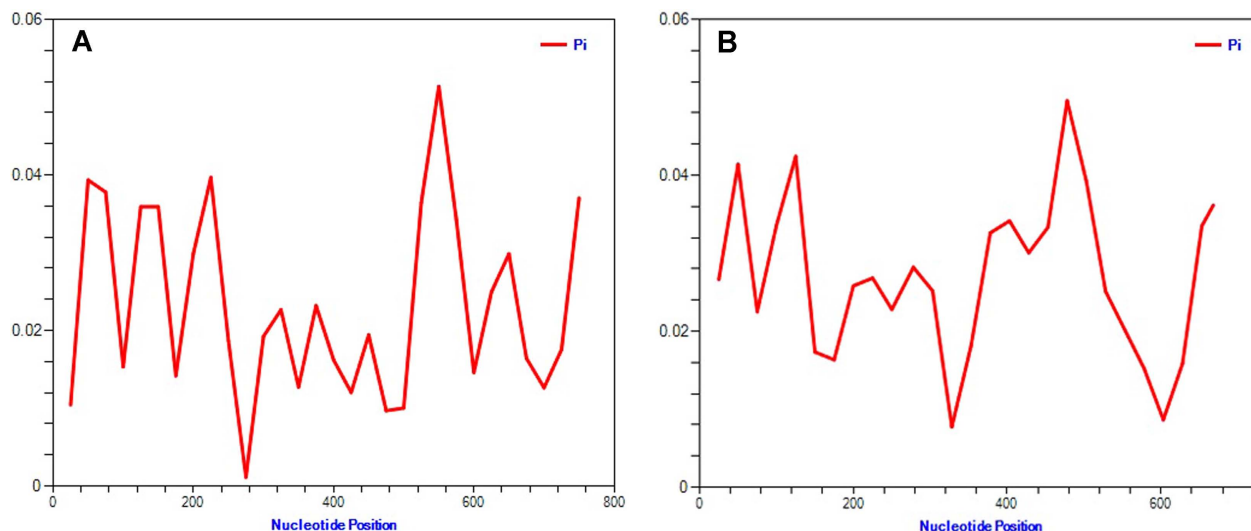


Fig. 8. Trend of polymorphism along the N and NSs genes in the TSWV populations, (A) N gene and (B) NSs gene comprising of 35 and 35 strains/isolates, respectively. Pi stands for nucleotide diversity. The curves were generated by sliding windows with 50 and 25 as the window and step sizes, respectively.

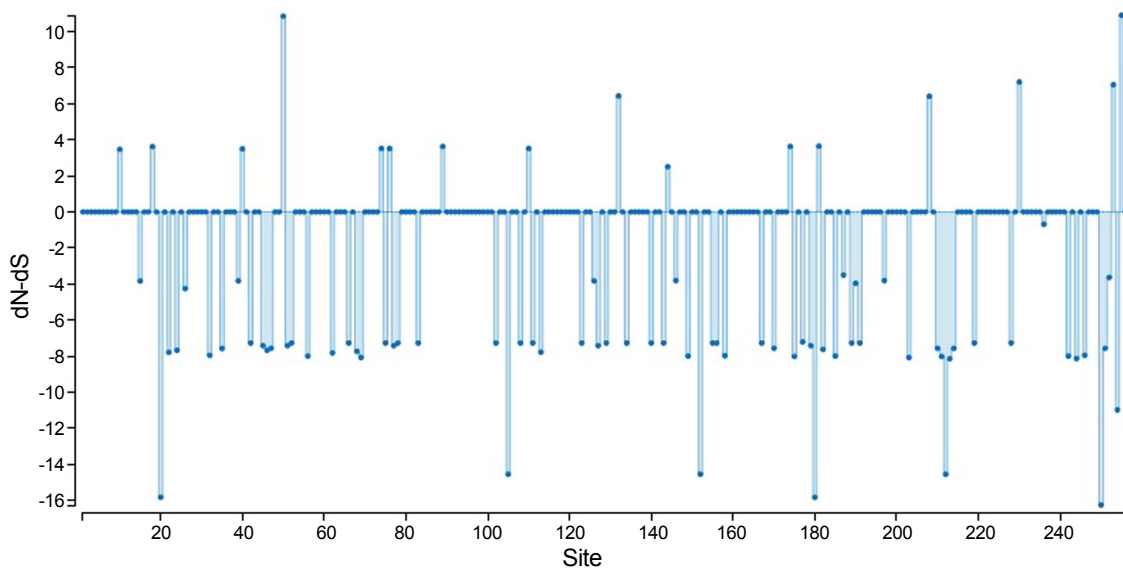


Fig. 9. SLAC site graph to identify positively and negatively codons/sites.

between Iranian populations was found when comparing the Iran and China populations. In both genes F_{ST} values were higher than 0.25 (Tables 8, 9). Between phylogroups, the maximum F_{ST} values were evaluated for clade I versus clade III (0.886) and clade III versus clade VII (0.909) for N and NSs genes, respectively, and the minimum F_{ST} values were obtained between clade III and clade V (0.666) (for N gene) and clade I and clade II (0.300) (for NSs gene) (Tables 8, 9).

Discussion

Viral diseases are one of the serious problems that reduce quality and quantity of tomato crops in Iran. In the present research, we observed a variety of viral symptoms including mild and sever mosaic, necrotic spots on the leaves, yellow spots on the fruits, malformations, and chlorosis in the different tomato fields of West and Northwest in Iran incurring significant losses. The study confirmed the

Table 6. Summary of demography test statistics between TSWV N gene populations

Comparisons	π	Tajima's D	Fu and Li's D*	Fu and Li's F*
All	0.02335	-0.74406 ns	-1.41358 ns	-1.40392 ns
Phylogroups				
Clade I	0.00346	-1.59698 ns	-1.67606 ns	-1.88212 ns
Clade II	0.00538	0.89518 ns	0.89518 ns	0.87188 ns
Clade III	0.00172	nd	nd	nd
Clade V	0.00646	-0.83379 ns	-0.83379 ns	-0.83370 ns
Clade VI	0.00517	-1.19267 ns	-1.19267 ns	-1.25779 ns
Clade VII	0.00726	-1.31354 ns	-1.40010 ns	-1.52101 ns
Geographic regions				
China	0.00726	-1.31354 ns	-1.40010 ns	-1.52101 ns
Iran	0.00538	0.89518 ns	0.89518 ns	0.87188 ns
South Korea	0.00517	-1.19267 ns	-1.19267 ns	-1.25779 ns
USA	0.01314	-0.69768 ns	-0.69768 ns	-0.72291 ns
Spain	0.01473	-1.13302 ns	-1.13302 ns	-1.22000 ns
Italy	0.00267	-0.31466 ns	-0.21471 ns	-0.25135 ns
France	0.00194	-0.75445 ns	-0.75445 ns	-0.67466 ns

$P > 0.10$, $0.10 > P > 0.05$

nd: Four or more sequences are needed to compute Tajima's and Fu and Li's statistics

π : Nucleotide diversity per site

ns not significant

Table 7. Summary of demography test statistics between TSWV NSs gene populations

Comparisons	π	Tajima's D	Fu and Li's D*	Fu and Li's F*
All	0.02651	-0.51290 ns	-1.31718 ns	-1.23457 ns
Phylogroups				
Clade I	0.00475	-1.83992*	-1.95542 ns	-2.19059 ns
Clade II	0.00978	1.11654 ns	1.11654 ns	1.12320 ns
Clade III	0.00176	-1.04849 ns	-1.04849 ns	-1.05189 ns
Clade V	0.00505	-0.85853 ns	-0.72792 ns	-0.83001 ns
Clade VI	0.00513	-0.81734 ns	-0.81734 ns	-0.79606 ns
Clade VII	0.00587	nd	nd	nd
Geographic regions				
China	0.00505	-0.85853 ns	-0.72792 ns	-0.83001 ns
Iran	0.00978	1.11654 ns	1.11654 ns	1.12320 ns
South Korea	0.00176	-1.04849 ns	-1.04849 ns	-1.05189 ns
USA	0.02346	-0.86239 ns	-0.86239 ns	-0.90711 ns
Spain	0.01877	-1.24792 ns	-1.24792*	-1.34590 ns
Italy	0.00450	-0.73528 ns	-0.68255 ns	-0.74634 ns
France	0.00318	-0.06501 ns	-0.06501 ns	-0.06004 ns

* $P < 0.05$

$0.10 > P > 0.05$

presence of TSWV in West and Northwest of Iran, even though there are some reports of TSWV from Iran based on symptoms and serological tests in tomato, soybean, potato, tobacco, peanut, and ornamental plants (Golnaraghi et al., 2001a, 2001b; Pourrahim et al., 2001). However, the genetic variation of TSWV isolates from tomato has not

been studied. On the other hand, the symptoms of TSWV on infected tomato leaves and fruits are similar to TYRV that was reported from Iran. It is possible that similarity of symptoms to TYRV is due to the cross reactions of different species in the same serogroup (Harrison, 2002; Pang et al., 1994) and therefore serological assays may not detect

Table 8. Summary of gene flow and genetic differentiation estimates between TSWV Nucleocapsid gene populations

Comparisons	^a K _S *	^a K _{ST} *	K _S *, K _{ST} * <i>p value</i>	^a Z*	<i>P value</i>	Snn	<i>P value</i>	^b F _{ST}
Phylogroups								
Clade I (n=12)/Clade II (n=4)	1.20693	0.30871	0.0010 **	3.37598	0.0000 ***	1.000	0.0020 **	0.729
Clade I (n=12)/Clade III (n=3)	1.11470	0.36319	0.0010 **	3.28415	0.0000 ***	1.000	0.0020 **	0.886
Clade I (n=12)/Clade V (n=4)	1.25705	0.34503	0.0010 **	3.41349	0.0020 **	1.000	0.0030 **	0.796
Clade I (n=12)/Clade VI (n=5)	1.23692	0.39751	0.0000 ***	3.44817	0.0000 ***	1.000	0.0000 ***	0.847
Clade I (n=12)/Clade VII (n=7)	1.33817	0.39906	0.0000 ***	3.57779	0.0000 ***	1.000	0.0000 ***	0.824
Clade II (n=4)/Clade III (n=3)	1.22876	0.47662	0.0290 *	1.43885	0.0260 *	1.000	0.0290 *	0.881
Clade II (n=4)/Clade V (n=4)	1.62731	0.36107	0.0330 *	1.67571	0.0330 *	1.000	0.0330 *	0.812
Clade II (n=4)/Clade VI (n=5)	1.50091	0.41024	0.0020 **	1.92111	0.0110 *	1.000	0.0110 *	0.856
Clade II (n=4)/Clade VII (n=7)	1.64246	0.35708	0.0030 **	2.38155	0.0010 **	1.000	0.0030 **	0.835
Clade III (n=3)/Clade V (n=4)	1.42925	0.26868	0.0270 *	1.43325	0.0200 *	1.000	0.0270 *	0.666
Clade III (n=3)/Clade VI (n=5)	1.32077	0.43240	0.0030 **	1.72670	0.0160 *	1.000	0.0160 *	0.885
Clade III (n=3)/Clade VII (n=7)	1.54595	0.34576	0.0040 **	2.24862	0.0020 **	1.000	1.000	0.860
Clade V (n=4)/Clade VI (n=5)	1.62121	0.36487	0.0020 **	1.93806	0.0110 *	1.000	0.0110 *	0.828
Clade V (n=4)/Clade VII (n=7)	1.72838	0.32373	0.0040 **	2.40607	0.0040 **	1.000	0.0040 **	0.808
Clade VI (n=5)/Clade VII (n=7)	1.63675	0.27381	0.0020 **	2.51815	0.0010 **	1.000	0.0020 **	0.691
Geographic regions								
China (n=7)/Iran (n=4)	1.64246	0.35708	0.0030 **	2.38155	0.0030 **	1.000	0.0030 **	0.835
China (n=7)/South Korea (n=5)	1.63675	0.27381	0.0000 ***	2.51815	0.0000 ***	1.000	0.0000 ***	0.691
China (n=7)/USA (n=4)	1.75307	0.30168	0.0030 **	2.39283	0.0020 **	1.000	0.0030 **	0.682
China (n=7)/France (n=4)	1.47644	0.37948	0.0020 **	2.35672	0.0010 **	1.000	0.0020 **	0.849
China (n=7)/Italy (n=6)	1.38525	0.41460	0.0000 ***	2.65988	0.0000 ***	1.000	0.0000 ***	0.835
China (n=7)/Spain (n=5)	1.93155	0.27321	0.0010 **	2.54973	0.0010 **	1.000	0.0010 **	0.684
Iran (n=4)/South Korea (n=5)	1.50091	0.41024	0.0010 **	1.92111	0.0070 **	1.000	0.0070 **	0.856
Iran (n=4)/USA (n=4)	1.67050	0.32079	0.0210 *	1.75615	0.0210 *	0.875	0.0520 ns	0.654
Iran (n=4)/France (n=4)	1.18641	0.39621	0.0240 *	1.67815	0.0240 *	1.000	0.0240 *	0.763
Iran (n=4)/Italy (n=6)	1.14630	0.39610	0.0010 **	2.21408	0.0000 ***	1.000	0.0220 *	0.748
Iran (n=4)/Spain (n=5)	1.97259	0.25276	0.0180 *	1.96706	0.0180 *	0.888	0.0310 *	0.659
South Korea (n=5)/USA (n=4)	1.65577	0.33821	0.0040 **	1.92297	0.0040 **	1.000	0.0040 **	0.694
South Korea (n=5)/France (n=4)	1.26849	0.45111	0.0000 ***	1.93333	0.0090 **	1.000	0.0090 **	0.874
South Korea (n=5)/Italy (n=6)	1.21066	0.46098	0.0000 ***	2.35725	0.0000 ***	1.000	0.0000 ***	0.860
South Korea (n=5)/Spain (n=5)	1.90996	0.28361	0.0060 **	2.12826	0.0060 **	0.980	0.0060 **	0.696
USA (n=4)/France (n=4)	1.37997	0.32198	0.0360 *	1.93816	0.0210 *	0.875	0.0360 *	0.583
USA (n=4)/Italy (n=6)	1.27534	0.33789	0.0140 *	2.43669	0.0160 *	0.900	0.0280 *	0.560
USA (n=4)/Spain (n=5)	2.12744	0.10406	0.0340 *	2.29517	0.0340 *	0.777	0.0590 ns	0.154
France (n=4)/Italy (n=6)	0.95261	-0.02929	0.5400 ns	2.96652	0.7700 ns	0.450	0.4900 ns	0.027
France (n=4)/Spain (n=5)	1.74016	0.25209	0.0140 *	2.15313	0.0130 *	0.888	0.0140 *	0.600
Italy (n=6)/ Spain (n=5)	1.54757	0.29211	0.0000 ***	2.53595	0.0000 ***	0.909	0.0070 **	0.576

ns, not significant

*0.01 < P < 0.05; **0.001 < P < 0.01; ***P < 0.001

^aK*, K_{ST}*, Z*, and Snn are test statistics of genetic differentiation^bF_{ST}, coefficient of gene differentiation, which measures inter-population diversity

Numbers in bold are maximum and minimum values between populations

TSWV properly. In this study we used a more sensitive method and sequences different genomic region of TSWV to confirm TSWV infection and also investigate the genetic variation of TSWV isolates from tomato plants.

As noted before, virus symptoms were observed in most tomato fields that were visited in different regions; however, necrotic spots on the leaves, common mosaic and yellowing were most common symptoms that observed in

Table 9. Summary of gene flow and genetic differentiation estimates between TSWV NSs gene populations

Comparisons	^a K _S *	^a K _{ST} *	K _S *, K _{ST} * <i>p value</i>	^a Z*	<i>P value</i>	Snn	<i>P value</i>	^b F _{ST}
Phylogroups								
Clade I (n=12)/Clade II (n=4)	1.46687	0.12186	0.0000 ***	3.49981	0.0000 ***	0.854	0.0200 *	0.300
Clade I (n=12)/Clade III (n=5)	1.20000	0.39658	0.0000 ***	3.27723	0.0000 ***	1.000	0.0000 ***	0.871
Clade I (n=12)/Clade V (n=7)	1.37205	0.39909	0.0000 ***	3.52921	0.0000 ***	1.000	0.0000 ***	0.863
Clade I (n=12)/Clade VI (n=4)	1.38060	0.34556	0.0010 **	3.37058	0.0000 ***	1.000	0.0010 **	0.860
Clade I (n=12)/Clade VII (n=3)	1.37565	0.32940	0.0020 **	3.31451	0.0020 **	1.000	0.0020 **	0.870
Clade II (n=4)/Clade III (n=5)	1.21189	0.44625	0.0010 **	1.96515	0.0070 **	1.000	0.0080 **	0.794
Clade II (n=4)/Clade V (n=7)	1.57717	0.35318	0.0030 **	2.45677	0.0020 **	1.000	0.0040 **	0.808
Clade II (n=4)/Clade VI (n=4)	1.75668	0.33468	0.0210 *	1.69769	0.0220 *	1.000	0.0300 *	0.806
Clade II (n=4)/Clade VII (n=3)	1.86390	0.32613	0.0290 *	1.42739	0.0180 *	1.000	0.0290 *	0.824
Clade III (n=5)/Clade V (n=7)	1.12971	0.34856	0.0000 ***	2.49338	0.0010 **	1.000	0.0020 **	0.726
Clade III (n=5)/Clade VI (n=4)	1.00484	0.55002	0.0090 **	1.96539	0.0090 **	1.000	0.0090 **	0.903
Clade III (n=5)/Clade VII (n=3)	0.89730	0.59632	0.0130 *	1.77873	0.0130 *	1.000	0.0130 *	0.909
Clade V (n=7)/Clade VI (n=4)	1.42928	0.41366	0.0020 **	2.41915	0.0030 **	1.000	0.0030 **	0.882
Clade V (n=7)/Clade VII (n=3)	1.42832	0.40692	0.0110 *	2.30258	0.0140 *	1.000	0.0140 *	0.888
Clade VI (n=4)/Clade VII (n=3)	1.51881	0.23044	0.0210 *	1.43903	0.0210 *	1.000	0.0210 *	0.594
Geographic regions								
China (n=7)/Iran (n=4)	1.57717	0.35318	0.0040 **	2.45677	0.0030 **	1.000	0.0040 **	0.808
China (n=7)/South Korea (n=5)	1.12971	0.34856	0.0020 **	2.49338	0.0020 **	1.000	0.0060 **	0.726
China (n=7)/USA (n=4)	1.70515	0.33467	0.0050 **	2.47380	0.0050 **	0.909	0.0130 *	0.686
China (n=7)/France (n=4)	1.31461	0.42712	0.0020 **	2.37166	0.0010 **	1.000	0.0020 **	0.883
China (n=7)/Italy (n=6)	1.36362	0.42559	0.0010 **	2.66747	0.0010 **	1.000	0.0010 **	0.867
China (n=7)/Spain (n=5)	1.71164	0.33597	0.0010 **	2.59190	0.0010 **	0.916	0.0070 **	0.715
Iran (n=4)/South Korea (n=5)	1.21189	0.44625	0.0020 **	1.96515	0.0080 **	1.000	0.0090 **	0.794
Iran (n=4)/USA (n=4)	2.23945	0.17655	0.0630 ns	1.94510	0.0760 ns	0.875	0.0760 ns	0.535
Iran (n=4)/France (n=4)	1.55602	0.13556	0.0210 *	1.99766	0.0370 *	0.750	0.0650 ns	0.320
Iran (n=4)/Italy (n=6)	1.54906	0.14538	0.0040 **	2.39122	0.0050 **	0.850	0.0140 *	0.320
Iran (n=4)/Spain (n=5)	2.14297	0.20121	0.0090 **	2.13111	0.0130 *	0.888	0.0240 *	0.574
South Korea (n=5)/USA (n=4)	1.39106	0.42220	0.0060 **	2.07020	0.0050 **	0.888	0.0260 *	0.664
South Korea (n=5)/France (n=4)	0.84431	0.57144	0.0070 **	1.96109	0.0070 **	1.000	0.0070 **	0.900
South Korea (n=5)/Italy (n=6)	1.04170	0.49558	0.0010 **	2.32213	0.0010 **	1.000	0.0010 **	0.876
South Korea (n=5)/Spain (n=5)	1.45206	0.39946	0.0020 **	2.30467	0.0030 **	0.900	0.0280 *	0.700
USA (n=4)/France (n=4)	1.77998	0.24882	0.0900 ns	2.12127	0.0850 ns	0.833	0.0900 ns	0.582
USA (n=4)/Italy (n=6)	1.69837	0.25247	0.0180 *	2.56118	0.0240 *	0.850	0.0440 *	0.549
USA (n=4)/Spain (n=5)	2.32215	0.06967	0.1290 ns	2.31178	0.0280 *	0.777	0.0640 ns	0.052
France (n=4)/Italy (n=6)	1.24275	0.02111	0.3070 ns	2.83881	0.3320 ns	0.550	0.3240 ns	0.102
France (n=4)/Spain (n=5)	1.77540	0.27052	0.0130 *	2.12125	0.0130 *	0.888	0.0210 *	0.630
Italy (n=6)/ Spain (n=5)	1.70676	0.27967	0.0020 **	2.50834	0.0030 **	0.909	0.0080 **	0.605

early summer; whereas yellow spot and ring spot symptoms were observed in middle of summer and autumn. Yellow spots on the fruit symptoms are the typical symptoms caused by TSWV. Different symptoms may show the presence of other viruses and/or mix infections by common viruses. There are other viruses that infected tomato including TSWV, CMV, *Tomato mosaic virus* (ToMV), TMV, *Potato virus X* (PVX), *Tomato ringspot virus* (ToRV), and

Beet curly top virus (BCTV) that have been reported in the most regions of tomato fields (Escriu et al., 2003; Hanssen and Lapidot, 2012).

In tomato plants showing a range of symptoms, mixed infection of CMV and ToMV was reported. Notably, for the geographical region that we collected tomato samples, a high rate of CMV infection in the cucurbits and tomato crops has reported (Sokhandan et al., 2008; Valizadeh

et al., 2011). Therefore, identifying TSWV based on the symptom is complex and there is a need to molecular methods for detection of TSWV in tomato plants.

The experimental host range assay provided further evidence for the presence of TSWV in the examined isolates because the inoculated plants developed symptoms resembling those of TSWV infected plants and showed positive reactions in the ELISA. However, considering the serological cross reaction between TSWV and other close viral species (Harrison, 2002; Pang et al., 1994;) we still need to confirm the presence of TSWV by a more sensitive and accurate method such as RT-PCR.

RT-PCR using specific primers to different regions of TSWV genome followed by sequencing was the most definitive approach for the identification of TSWV. The degenerate primers designed based on the consensus nucleotide sequences of RdRp initially provided the identification to the genus level (Tsompana et al., 2005). In contrast, the sequence of N genes form Orthotospoviruses are more diverse, thus it can be the key gene target for classification of a *Orthotospovirus* (Margaria et al., 2015) and it could be reliable to differentiate the *Orthotospovirus* members because the consensus nucleotide sequences of N genes in *Orthotospovirus*. The samples that not yield any amplification in RT-PCR raised the possibility that other viruses could be present in those plants. Whether negative samples in this study infected by other viruses need to be analyzed further. Further testing showed the presence of TMV and CMV in the remaining samples (data not shown). Presence of other viruses such as CMV in tomato fields is not unusual in the studied area (Bashir et al., 2006; Sokhandan et al., 2008).

Phylogenetic analysis showed that the new Iranian isolates were grouped into a distinct cluster. The isolates from Italy and France were placed close to the Iranian isolates. Such a phylogenetic relation between the isolates may be related to the virus transmission. In addition of the polymerase gene, some studies have focused on the diversity of N and NSs genes for strain identification and genetic analysis (Chiemsombat et al., 2008; Nischwitz et al., 2007; Pappu et al., 2009; Zheng et al., 2008).

Sequence analysis for N and NSs genes revealed that the Iranian isolates were grouped with several isolates from Europe. The Iranian isolates were differentiated from European isolates by high bootstrap, but separated from South Korean, Japanese, and American isolates. It is possible that origin of Iranian TSWV isolates is different from other Asian countries and they could have been introduced from Europe. As a result, TSWV isolates were not grouped according to geographical origin, as previously reported by

(Zindovic et al., 2014), whereas most Asian isolates (South Korea, Japan, and China) were grouped in a distinct cluster. Global trade in plant products might be the main contributing factor for TSWV spread throughout the world.

In addition, the high nucleotide identity and similarities among the Iranian isolates based on the L, N, and NSs genes showed that minor evolutionary differences exist amongst the isolates. However, the minor isolate differences are evidence of some mutations taking place in different genomic regions. Population genetics studies should be conducted to find the systematic and random forces of evolution factors on the new isolates.

It is clear that mutation has a significant effect in shaping the population structure of TSWV (Kaye et al., 2011). Also, survey of the variability in the genetic structure of plant virus populations is crucial for a better understanding of virus evolution and interaction between plant and virus (Tsompana et al., 2005). Hence, detection and sequencing of TSWV isolates have an important implication for crop disease management.

Because TSWV has numerous opportunities to exchange genetic information with other strain or viruses (Kaye et al., 2011), the complete genome sequences of TSWV isolates with molecular diversity will provide fundamental information to elucidate molecular interactions between host plants, TSWV, and geographic regions. Furthermore, this information can be applied practically in molecular breeding to develop TSWV-resistant vegetables and crops.

In this research, genetic diversity and molecular evolutionary were examined to identify haplotypes and infer genetic relationships between populations of the TSWV based on the N and NSs genes. Comparison of nucleotide diversity (π) showed that the lowest nucleotide diversity (0.002) was found for clade III in comparison with other clades and among geographical populations was obtained for French TSWV population (0.002) based on the N gene. Total of nucleotide sequence diversity was 0.064. One of the most important reasons that may have caused the nucleotide diversity of this phylogroup and France population to be less than other phylogroups is due to strong negative selection ($\omega=0$) (Yu et al., 2010). The largest dN/dS ratios were found for Iran population (0.484) and clade II (0.484) in the N gene which isolated from different varieties of *Solanum lycopersicum*, and based on the NSs gene, the maximum dN/dS ratios were obtained for China population (0.329) and clade V (0.0329) from different hosts (Table 3). The obtained ω ratios were less than 1 for all populations, suggesting that the negative selection has been occurred for all the TSWV populations. Some studies on viral N genes in other RNA viruses have revealed that this gene can be

conserved under negative selection (Gao et al., 2016).

Haplotype diversity (Hd) measures can range from 0, meaning no diversity, to 1.000, which indicates high levels of haplotype diversity (Nei and Tajima, 1981). Our data showed that Hd was 0.983 to 0.990, indicating very high levels of diversity for each locus. These levels were similar to those indicated for global populations, which were 0.857 to 1 (Tsompana et al., 2005). Most studies of *Orthotospovirus* spp. have focused on the genetic diversity of the N gene for strain identification or the NSs gene (Chen et al., 2006; Pappu et al., 2006). In addition, these studies should include noncoding regions of the TSWV genome to provide enhanced statistical analysis of the data (Rosenberg and Nordborg, 2002).

Nucleotide diversity analysis by “Sliding window” option with 50 and 25 as the window and step size, respectively, revealed high polymorphism at middle regions and 5' parts of both N and NSs genes (Figs. 8A, 8B, Supplementary). This information can help to avoid designing specific primers for these regions of viral genome for virus detection.

Non-significantly negative and positive values in all clades were calculated by neutrality test statics (Tajima's D, Fu and Li's D* and F*) in both genes except significantly negative values for clade I in the NSs gene. Between geographic populations, Fu and Li's D* test static was only significantly negative for Spanish population in the NSs gene which shows low-frequency polymorphism and only Iranian population was non-significantly positive in both N and NSs genes (Tables 4, 5). In geographical regions, the maximum F_{ST} values between South Korea and France populations based on the N ($F_{ST} = 0.874$) and NSs ($F_{ST} = 0.900$) genes could be due to long distances between these countries. The results obtained from phylogenetic analysis and genetic differentiation were in line with each together. For instance, the minimum F_{ST} value was obtained between USA/Spain populations ($F_{ST} = 0.052$) based on the NSs gene, and France/Italy populations ($F_{ST} = 0.027$) based on the N gene which they clustered together in a sub-clade in phylogenetic tree (Figs. 5, 6).

Using SLAC algorithm which implemented in Datamonkey server within HyPhy software package, negative selection pressure was observed at position 258 codon and it seems that negative pressure is one of the most important reasons for survival and evolution of the TSWV. In this study and several studies the N gene was under negative pressure (Timmerman-Vaughan et al., 2014).

In summary, the first detailed molecular characterization followed by genetic diversity and population structure studies of TSWV from tomato in Iran has shed light on the

nature of TSWV populations in tomato and provides a basis for further studies on TSWV in other crops and toward developing a sound disease management program.

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgments

The research was supported by University of Zanjan. The Authors thank Dr. Hanu Pappu for review the paper and Dr. Afshin Hasani-Mehraban for TSWV antibody.

References

- Bananej, K., Ahoonmanesh, A., Shahraeen, N. and Lesemann, D. 1996. Identification of *Tomato spotted wilt virus* from tomato fields in Varamin area. *Iran. J. Plant Pathol.* 32:29-30.
- Bashir, N. S., Kalhor, M. R. and Zarghani, S. N. 2006. Detection, differentiation and phylogenetic analysis of *Cucumber mosaic virus* isolates from cucurbits in the northwest region of Iran. *Virus Genes* 32:277-288.
- Brittlebank, C. 1919. Tomato diseases. *J. Agric. Victoria* 17:231-235.
- Chen, T. C., Huang, C. W., Kuo, Y. W., Liu, F. L., Yuan, C. H., Hsu, H. T. and Yeh, S. D. 2006. Identification of common epitopes on a conserved region of NSs proteins among tospoviruses of *Watermelon silver mottle virus* serogroup. *Phytopathology* 96:1296-1304.
- Chen, T. C., Tsai, W. T., Kang, Y. C., Wang, Y. C. and Yeh, S. D. 2016. Using monoclonal antibodies against the common epitopes of NSs proteins for the prompt detection and differentiation of tospoviruses prevalent in Euro-America and Asia regions. *Eur. J. Plant Pathol.* 144:509-524.
- Chiemsoombat, P., Gajanandana, O., Warin, N., Hongprayoon, R., Bhunchoth, A. and Pongsapich, P. 2008. Biological and molecular characterization of tospoviruses in Thailand. *Arch. Virol.* 153:571-577.
- Chu, F. H., Chao, C. H., Chung, M. H., Chen, C. C. and Yeh, S. D. 2001. Completion of the genome sequence of *Watermelon silver mottle virus* and utilization of degenerate primers for detecting tospoviruses in five serogroups. *Phytopathology* 91:361-368.
- Chung, C. T., Niemela, S. L. and Miller, R. H. 1989. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A.* 86:2172-2175.
- Escriu, F., Fraile, A. and Garcia-Arenal, F. 2003. The evolution of virulence in a plant virus. *Evolution* 57:755-765.
- 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 *Arabidopsis mosaic virus* based on the CP gene sequence. *Arch. Virol.* 161:1047-1051.
- Ghotbi, T. and Shahraeen, N. 2012. Incidence and distribution of viruses infecting propagated ornamentals in Northern Iran. *Int. Res. J. Microbiol.* 3:373-381.
- Golnaraghi, A. R., Shahraeen, N., Pourrahim, R., Ghorbani, S. and Farzadfar, S. 2001a. First report of a Tosspovirus infection of peanuts in Iran. *Plant Dis.* 85:1286.
- Golnaraghi, A. R., Shahraeen, N., Pourrahim, R., Ghorbani, S. and Farzadfar, S. 2001b. First report of *Tomato spotted wilt virus* on soybean in Iran. *Plant Dis.* 85:1290.
- 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.
- Golnaraghi, A. R., Pourrahim, R., Farzadfar, S. and Ahoonmanesh, A. 2007. Identification and partial characterization of a Tosspovirus causing leaf and stem necrosis on potato. *Plant Pathol. J.* 6:227-234.
- Hajiabadi, A. M., Asaei, F., Abdollahi Mandoulakani, B. and Rastgou, M. 2012. Natural incidence of tomato viruses in the North of Iran. *Phytopathol. Mediterr.* 51:390-396.
- Hanssen, I. M. and Lapidot, M. 2012. Major tomato viruses in the Mediterranean Basin. *Adv. Virus Res.* 84:31-66.
- Harrison, B. D. 2002. Virus variation in relation to resistance-breaking in plants. *Euphytica* 124:181-192.
- Hudson, R. R. 2000. A new statistic for detecting genetic differentiation. *Genetics* 155:2011-2014.
- Hudson, R. R., Boos, D. D. and Kaplan, N. L. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* 9:138-151.
- Jones, D. R. 2005. Plant viruses transmitted by thrips. *Eur. J. Plant Pathol.* 113:119-157.
- Kaye, A. C., Moyer, J. W., Parks, E. J., Carbone, I. and Cubeta, M. A. 2011. Population genetic analysis of *Tomato spotted wilt virus* on peanut in North Carolina and Virginia. *Phytopathology* 101:147-153.
- Kimura, S. and Sinha, N. 2008. Tomato (*Solanum lycopersicum*): a model fruit-bearing crop. *Cold Spring Harb. Protoc.* 2008:pdb.emo105.
- King, A. M. Q., Adams, M. J., Carstens, E. B. and Lefkowitz, E. J. 2011. Virus taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses. 9th ed. Elsevier, Cambridge, MA, USA, 1327 pp.
- Lian, S., Lee, J. S., Cho, W. K., Yu, J., Kim, M. K., Choi, H. S. and Kim, K. H. 2013. Phylogenetic and recombination analysis of *Tomato spotted wilt virus*. *PLoS One* 8:e63380.
- Margaria, P., Miozzi, L., Ciuffo, M., Pappu, H. and Turina, M. 2015. The first complete genome sequences of two distinct European *Tomato spotted wilt virus* isolates. *Arch. Virol.* 160:591-595.
- Muhire, B. M., Varsani, A. and Martin, D. P. 2014. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS One* 9:e108277.
- Nischwitz, C., Pappu, H., Mullis, S., Sparks, A., Langston, D., Csinos, A. and Gitaitis, R. 2007. Phylogenetic analysis of *Iris yellow spot virus* isolates from onion (*Allium cepa*) in Georgia (USA) and Peru. *J. Phytopathol.* 155:531-535.
- Pang, S. Z., Bock, J. H., Gonsalves, C., Slightom, J. L. and Gonsalves, D. 1994. Resistance of transgenic *Nicotiana benthamiana* plants to *Tomato spotted wilt* and *Impatiens necrotic spot tospoviruses*: evidence of involvement of the N protein and N gene RNA in resistance. *Phytopathology* 84:243-249.
- Pappu, H., Du Toit, L., Schwartz, H. and Mohan, S. 2006. Sequence diversity of the nucleoprotein gene of *Iris yellow spot virus* (genus *Tospovirus*, family *Bunyaviridae*) isolates from the western region of the United States. *Arch. Virol.* 151:1015-1023.
- Pappu, H., Jones, R. and Jain, R. 2009. Global status of tospovirus epidemics in diverse cropping systems: successes achieved and challenges ahead. *Virus Res.* 141:219-236.
- Parrella, G., Gognalons, P., Gebre-Selassie, K., Vovlas, C. and Marchoux, G. 2003. An update of the host range of *Tomato spotted wilt virus*. *J. Plant Pathol.* 85:227-264.
- Pond, S. L. K. and Muse, S. V. 2005. HyPhy: hypothesis testing using phylogenies. In: *Statistical methods in molecular evolution*, pp. 125-181. Springer, New York, USA.
- Pourrahim, R., Farzadfar, S., Moini, A., Shahraeen, N. and Ahoonmanesh, A. 2001. First report of *Tomato spotted wilt virus* on potatoes in Iran. *Plant Dis.* 85:442.
- Rosenberg, N. A. and Nordborg, M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat. Rev. Genet.* 3:380-390.
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E. and Sanchez-Garcia, A. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* 34:3299-3302.
- Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.
- Samuel, G., Bald, J. G. and Pittman, H. A. 1930. Investigations on "Spotted Wilt" of Tomatoes. H. J. Green, Government Printer, Melbourne, Australia. 120 pp.
- Sherwood, J., German, T., Moyer, J. and Ullman, D. 2003. Tomato spotted wilt. *The Plant Health Instructor*. doi:10.1094/PHI-1-2003-0613-02.
- Sokhandan, B. N., Nematollahi, S. and Torabi, E. 2008. *Cucumber mosaic virus* subgroup IA frequently occurs in the north-west Iran. *Acta Virol.* 52:237-242.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- Takeda, A., Sugiyama, K., Nagano, H., Mori, M., Kaido, M., Mise, K., Tsuda, S. and Okuno, T. 2002. Identification of a novel RNA silencing suppressor, NSs protein of *Tomato spotted wilt virus*. *FEBS Lett.* 532:75-79.
- Timmerman-Vaughan, G. M., Lister, R., Cooper, R. and Tang, J. 2014. Phylogenetic analysis of New Zealand *Tomato spot-*

- ted wilt virus* isolates suggests likely incursion history scenarios and mechanisms for population evolution. *Arch. Virol.* 159:993-1003.
- Tripathi, D., Raikhy, G. and Pappu, H. R. 2015. Movement and nucleocapsid proteins coded by two tospovirus species interact through multiple binding regions in mixed infections. *Virology* 478:137-147.
- Tsompana, M., Abad, J., Purugganan, M. and Moyer, J. 2005. The molecular population genetics of the *Tomato spotted wilt virus* (TSWV) genome. *Mol. Ecol.* 14:53-66.
- Valizadeh, M., Valizadeh, J. and Jafari, M. 2011. Identification, distribution and incidence of important tomato and cucurbits viruses in the southeast of Iran. *Am. J. Plant Physiol.* 6:242-251.
- Whitfield, A. E., Ullman, D. E. and German, T. L. 2005. Tospovirus-thrips interactions. *Annu. Rev. Phytopathol.* 43:459-489.
- Zhang, Z., Wang, D., Yu, C., Wang, Z., Dong, J., Shi, K. and Yuan, X. 2016. Identification of three new isolates of *Tomato spotted wilt virus* from different hosts in China: molecular diversity, phylogenetic and recombination analyses. *Viol. J.* 13:8.
- Zheng, Y. X., Chen, C. C., Yang, C. J., Yeh, S. D. and Jan, F. J. 2008. Identification and characterization of a tospovirus causing chlorotic ringspots on Phalaenopsis orchids. *Eur. J. Plant Pathol.* 120:199-209.
- Zindović, J., Ciuffo, M. and Turina, M. 2014. Molecular characterization of *Tomato spotted wilt virus* in Montenegro. *J. Plant Pathol.* 96:201-205.