INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
20–1236/2021/25–5–1061–1065
DOI: 10.17957/IJAB/15.1764
http://www.fspublishers.org

Short Communication

Molecular Analysis of *Tomato Spotted Wilt Virus* N Gene from Pakistan

Adnan Ahmad, Muhammad Naveed Aslam and Waqas Ashraf*
Department of Plant Pathology, Faculty of Agriculture and Environment, The Islamia University of Bahawalpur, Box. 63100, Bahawalpur, Pakistan
*For correspondence: waqasashraf@iub.edu.pk
Received 29 August 2020; Accepted 08 February 2021; Published 16 April 2021

Abstract

The molecular diversity of *Tomato spotted wilt virus* (TSWV) isolates from Pakistan was investigated based on nucleocapsid protein (N) gene. It revealed that the Pakistani isolates shared 98.2–99.3% (nt) and 95.7–97.4% (aa) identities with each other and 97.2–99.3% (nt) and 92.7–97.8% (aa) with other TSWV isolates retrieved from Genbank. In a neighbour-joining phylogenetic reconstruction, four Pakistani isolates clustered with South Korean and Turkish isolates and one with Australian and South African isolates. All the Pakistani isolates shared maximum genetic diversity (0.024–0.039 and 0.026–0.029) with South African (AJ296600) and Australian (AY879109) isolates and the lowest values of genetic diversity (0.009–0.013) was recorded with Czech Republican isolate (AJ296599). The gene differentiation co-efficient value (Fst) was 0.33 less than standard value noticed as 0.05145, between globally reported and Pakistani TSWV isolates considering it a frequent gene flow. The significant values of statistical tests based on three genetic differentiation analysis; K$\overset{*}{s}$; 4.86231, Z; 34.27778 and Snn; 0.55556 were observed. The negative values were recorded according to statistical test as Fu, & Li's D*(-2.55465), Fu, & Li's F*(-2.78103) and one positive value from Tajima's D (2.10697). These results indicate low polymorphism frequency in TSWV population of Pakistan. © 2021 Friends Science Publishers

Keywords: Molecular diversity; *Tomato spotted wilt virus*; Tospovirus

Introduction

Tomato spotted wilt virus (TSWV; type member of genus *Tospovirus*) is transmitted by thrips belonging to *Dictyothripsbetae Uzel, Thrips tabaci Lindeman, F. intonsatrybom* and *Frankliniella occidentalis* (most efficiently) infecting tomato around the globe (Ciuffo *et al.* 2008). The quasi-spherical viral particles of TSWV range from 80 to 120 nm in diameter. The capsid is surrounded by a lipid bilayer made of glycoprotein, which consists of + ssRNA tripartite genome having small (S), medium (M) and large (L) segments. Small segment contains 2900 nucleotides which encodes non-structural protein silencing suppressor (NSs) and nucleocapsid protein (Takeda *et al.* 2002). The medium segment having 4800 nucleotides encodes precursor glycoprotein (Gn-GC) and movement protein (NSm). The large segment comprises 8900 nucleotides which encodes RNA dependent RNA polymerase enzyme (RdRp) involved in replication of the virus (Soellick *et al.* 2000).

It is reported that TSWV infects more than 1300 plant species belonging to 92 families of mono and dicotyledonous plants causing a loss of more than 1 billion USD annually (Parrella *et al.* 2003). Typical symptoms caused by TSWV includes stunting, necrosis, chlorosis, bronzing and ringspots (Adkins 2000). Along with other pathogens (fungi and bacteria), the management of viral diseases is of supreme importance as the crop is infected by at least 136 characterized viral species (Xu *et al.* 2017). The amplified international trade and the global climatic variations are causing more frequent emergence of new viruses infecting tomato and other crops, while the previously reported viruses are becoming epidemics (Hanssen *et al.* 2010). It is critical to understand the diversity and evolution of plant viral pathogens for their management.

The molecular evolution study of plant RNA viruses is very crucial to understand the parameters of virus managements like geographical distribution, adaptation, and evolutionary process (Lauring and Andino 2010). The information of Pakistani native plant viruses regarding molecular characterization is not sufficient because of unknown genetic variability. Due to extensive tomato cultivation over a large area, a wide range of tomato infecting genetically diverse viral strains and species may exist in Pakistan and their genetic exchange can cause the evolution of new viral species. The current research work was conducted to explore the molecular variability of Pakistani TSWV- isolates. The findings of this study will be helpful in development of sustainable management

strategies for TSWV disease that will ultimately boost the quality of tomato production in Pakistan.

Materials and Methods

Samples collection

Symptomatic tomato leaves were collected from plants exhibiting foliar ring spots, stunting, necrosis, chlorosis, line pattern and bronzing associated with *Tomato spotted wilt virus*. These samples were collected from farmer's fields during April to August 2018 survey from Bahawalpur, Faisalabad, Lahore and Multan. Symptomatic and healthy leaf samples were tested for TSWV infection by using TSWV-specific-Bioreba Agri Strip ELISA Assay.

RNA extraction and Reverse transcriptase polymerase chain reaction (RT-PCR)

The total RNA content of ELISA positive plant leaves was isolated by using TRIzol® Reagent (Life Technologies, Carlsbad, U.S.A.) as per manufacturer's protocol. The isolated RNA was directly used in RT-PCR. TSWV-nucleocapsid protein (N-gene) was amplified by using primers pair of CP-5Bam and CP3-Pst (Antignus *et al.* 1997) in one-step RT-PCR. The reaction mixture was incubated at 42°C for 60 min for the synthesis of genomic cDNA of TSWV. The initial denaturation temperature was set at 95°C for 5 min and 35 cycles were set for template amplification for 60 s at 94°C, 60 s at 58°C than extension for 60 s at 72°C and final extension was done at 72°C for 10 min. The amplified PCR products were separated in 1% agarose gel electrophoresis. The positive PCR products were cleaned by using PCR purification kit (Quiagen) QIAquick® and sequenced in opposite directions.

Sequence analysis

The Transeq program (EMBOSS) was used to translate the obtained nucleotide sequences (Rice *et al.* 2000). The BLAST analysis of protein and nucleotide sequences was done to identify the TSWV. Closely related sequences were downloaded from NCBI and aligned by using CLUSTAL W program (Larkin *et al.* 2007). Also, the percent identities of amino acid and nucleotide sequences were recorded by BioEdit v. 7.2.6.1. The software MEGA 6 with default parameters and 1000 bootstrap replications was used to construct the phylogenetic tree and measure the evolutionary distance using neighbour joining method (Tamura *et al.* 2013).

Recombination analysis and selection pressure

The software DnaSP (v. 5.10) was explored to detect the deletion and insertion of nucleotides within studied sequences (Balasubramanian and Selvarajan 2014). DnaSP version 5.0 was used to evaluate the molecular diversity pattern at segregating sites by statistical analysis like diversity of nucleotides at all sites, Fu, & Li's F*, Fu, & Li's D*and Tajima's D (Librado and Rozas 2009).

Results

Samples collection

During field surveys, the TSWV infected samples were showing necrosis, chlorosis and bronzing on foliar parts. These symptoms are associated with TSWV infection.

PCR amplification and analysis of sequences

The primer pair CP-Pst and CP5-Bam amplified the ~800 DNA fragments in each positive sample. The obtained amplified PCR fragment was consisted of 777 nucleotides. The BLASTn analysis showed that nucleocapsid protein gene (N-gene) is a part of small (S) fragment and located between 152nd and 928th nucleotides of TSWV genome. Translate tool translated the coding sequence into 259 amino acids. Few variations were observed by using multiple sequence alignment of amino acid sequences of Pakistani TSWV isolates. The central region of all isolates was observed to be conserved (Fig. 1). The variations were in the 5' and 3' ends of N gene sequences. The nucleotide sequences of five TSWV isolates (TSWV-PK1-5) were carefully analysed and submitted into Genbank. Each sequence consisted of Uracil contents 33–34%, Guanine 18–19%, Cytosine 27–28% and Adenine 21–22% (Table 1). In BLAST analysis the nucleotides and amino acid sequences of all Pakistani-TSWV isolates showed 98.2–99.3% and 95.7–97.4% similarity with each other, respectively. While they showed 97.2–99.3% and 92.7–97.8% nucleotide and amino acid sequence similarity with previously reported TSWV isolates respectively. Highest percentage identities were observed with TSWV-AJ295699 (Czech Republic) and TSWV-AY744478 (USA) isolates.

Phylogenetic reconstruction

The phylogenetic tree was constructed among the sequences of nucleocapsid (N) gene sequences of new TSWV Pakistani isolates (TSWV-PK1-5) and others previously worldwide reported twenty-two TSWV-isolates through neighbour-joining. In phylogenetic tree twenty-seven TSWV isolates making two major groups. Four Pakistani TSWV isolates (TSWV1-2, 3 and 5) clustered with South Korean and Turkish isolates in group-II. While one of the Pakistani TSWV isolate (TSWV-4) clustered with Australian and South African isolates in Group-I (Fig. 2). Five American isolates (KU179581, KU179561, KU179591, KU179577 and KU179513) in group-II appeared to be rotted by Pakistani TSWV-isolates. The remaining globally reported TSWV- isolates clustered in another subgroup separately.

Table 1: Characteristics of nucleotide sequences of Pakistani TSWV isolates

Isolate	Accession No	(nt)	(aa)	A (%)	C (%)	G (%)	U (%)
TSWV-PK1	MN966565	777	259	34	20	24	22
TSWV-PK2	MN966566	777	259	33	20	24	23
TSWV-PK3	MN966567	777	259	34	20	23	23
TSWV-PK4	MN966568	777	259	33	20	24	23
TSWV-PK5	MN966569	777	259	33	20	24	23

* Number of nucleotides (nt), encoded amino acids (aa) and percentages of nucleotide bases *i.e.*, A (Adenine), C (Cytosine), G (Guanine) and U (Uracil)

Table 2: Genetic and molecular diversity of Pakistani TSWV isolates

Statistical test	Values
Evolutionary distance	0.007-0.219
Tajima's D	2.10697
Fu, & Li's D*	-2.55465
Fu, & Li's F*	-2.78103
Ks*	4.86231
Z	34.27778
Snn	0.55556
Fst	0.05145

Recombination analysis and selection pressure

The investigations of Single Nucleotide Polymorphism (SNP) and Insertions and Deletions (INDEL) in nucleotide sequences were done to confirm the nature of polymorphism among TSWV isolates. The INDELs were absent in studied sequences. Pakistani isolates were holding 0.007–0.219 evolutionary distance among themselves and 0.097–0.161 with isolates reported from other countries (Table 2). The highest evolutionary distance values were observed with isolates AY879109 (0.026–0.029) and AJ296600 (0.024–0.039) and lowest with isolate AJ296599 (0.009–0.013).

Genetic differentiation and gene flow analysis

In gene flow analysis, the gene differentiation coefficient value (Fst) was observed as 0.05145, which is less than standard value 0.33 (Table 2). This recorded value between world and Pakistani TSWV isolates show a frequent gene flow among them.

The genetic differentiation analysis was observed on the basis of three permutation statistical test as Snn; 0.55556, Z;34.27778and Ks*; 4.86231; (Table 2) (Balasubramanian and Selvarajan 2014).

Haplotype and nucleotide diversity

Haplotype and nucleotide diversity from all sites were performed by using statistical test like Fu, & Li's F*, Fu, & Li's D * and Tajima's D. These statistical tests were performed to determine molecular diversity pattern at segregation sites. The values recorded in these statistical tests were found negative *i.e.,* Fu, & Li's D*(-2.55465), Fu, & Li's F*(-2.78103) and one positive result from Tajima's D (2.10697) (Table 2). These negative results indicate low polymorphism frequency in Pakistani *Tomato spotted wilt virus* (TSWV) population (Tsompana *et al.* 2005).

Discussion

In the present study, the N gene of TSWV was selected for the characterization, because the nucleocapsid gene is considered as vital in classification of plant viruses as it encodes for the synthesis of capsid protein. Moreover, the interaction of plant viruses with their hosts and their life cycle are also dependent on capsid protein gene (Moury and Simon 2011). The reports of International Committee of Taxonomy of Viruses have suggested that, the nucleocpasid genes of plant viruses belonging to same specie hold < 80% nucleotide and amino acid sequences similarities (Ismaeil *et al.* 2015). In current study the amino acid and nucleotide sequences of all *Tomato spotted wilt virus* Pakistani isolates were 95.7–97.4% and 98.2–99.3% similar to each other, which mean that all of the isolates were in close association with each other as compared to world isolates. Also, the Pakistani isolates clustered in same group in phylogenetic reconstruction. These findings were also supported by the detection of frequent gene flow among Pakistani isolates. Moreover, the TSWV population was observed to be under purifying or negative selection as neutrality test like Fu, & Li's F*, Fu, & Li's D*and Tajima's D, showed negative values. Evolutionary distance describes the divergence of homologous sequences from their common ancestors (Rosenberg 2005). The lower genetic distance was observed among Pakistani isolates, which also strengthens the close association of Pakistani TSWV isolates.

Conclusion

The present study concludes that there exists a close association among Pakistani isolates and comprehensive surveys for detection and molecular characterization based on complete genomic sequence of the studied virus from different crops is recommended. This will be helpful to develop sustainable and comprehensive strategies for the management of TSWV in future.

Author Contributions

The author AA, MNA and WA, contributed to designing and executing lab work and basic writeup.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability

The data will be made avaialble on acceptable requests to the corresponding author.

```
TSWV-PK-2  SFSIRIAGAKYSSILFPFLHLIFIHFKCFAFQHSANFXLRLLWCHTSLGQSRGLCILHPD 60
TSWV-PK-1  SFSIRIAGAEYSSILFSFLHLIFIHFECFAFQHSANFXLRLLWCHTSLGQSRGLCILHPD 60
TSWV-PK-5  SFSIRIAGAEYSSILFSFLHLIFIHFKCFAFQHSANFXLRLLWCHTSLGQSRGLCILHPD 60
TSWV-PK-3  SFSIRIAGAEYSSILFPFLHLIFIHFKCFAFQHSANFXLRLPWCHTSLGQSRGLCILHPD 60
TSWV-PK-4  SFSIRIAGAKYSSILFSFLHLIFIHFKCFAFQHSANFXLRLPWCHTSLGQSRGLCILHPD 60
           *********:****** *********:********** ******************

TSWV-PK-2  I*PRQH*SSQSYQLKQ*EVSYLPALWQVSQTLHHQEVIHRLGSKGGKQS*I**Y*DSQNS 113
TSWV-PK-1  I*PRQH*SSQSYQLKQ*EVSYFPALWQASQTLHHQEVIHRLESKGGKQS*I**Y*DSQNS 113
TSWV-PK-5  I*PRQH*SSQGYQLKQ*EVSYLPALWQASQTLHHQEVTHRLESKGGKQS*I**Y*DSQNS 113
TSWV-PK-3  I*PRQH*SSQSYQLKQ*EVSYLPALWQASQTLHHQEVTHRLESKGGKQS*I**Y*DSQNS 113
TSWV-PK-4  I*PRQH*SSQSYQLKQ*EVSYLPALWQASQTLHHQEVIHRLESKGGKQS*I**Y*DSQNS 113
           **********.**********:*****.********* *** ******************

TSWV-PK-2  QFPLQA*P*SSYQAF*RSCQWLQSCLKFSLW*FYQK*NHFA**PSLYSDDSSGMSDMK*R 165
TSWV-PK-1  QFPLQA*P*SSYQAF*RSCQWLQSCLKFSLW*FYQK*NHFA**PSLYSDDSSGMSDMK*C 165
TSWV-PK-5  QFPLQA*P*SSYQAF*RSCQWLQSCLKFSLW*FYQK*NHFA**PSLYSDDSSGMSDMK*R 165
TSWV-PK-3  QFPLQA*P*SSYQAF*RSCQWLQSCLKFSLW*FYQK*NHFA**PSLYSDDSSGMSDMK*R 165
TSWV-PK-4  QFPLRA*P*SSYQAF*RSCQWLQSCLKFSLW*FYQK*NHFA**PSLCSDDSSGMSDMK*R 165
           ****:**************************************** ************

TSWV-PK-2  SSS*SGRGFPDKKS*S*MLPDSDLPQTQGLCLVSTKQQCFP**A*P*T**L*KLLYALTX 215
TSWV-PK-1  SSS*SGRGFPDKKS*S*MLPDSDLPQTQGLCLVSTKQQCFP**A*P*T**S*KLLYALTX 215
TSWV-PK-5  SSS*SGQGFPDKKS*S*MLPDSDLPQTQGLCLVSTKQQCFP**A*P*T**L*KLLYALTX 215
TSWV-PK-3  SSS*SGQGFPDKKS*S*MLPDSDLPQTQGLCLVSTKQQCFP**A*P*T**S*KLLYALTX 215
TSWV-PK-4  SSS*SGRGFPDKKS*S*MLPDSDLPQTQGLCLVSTKQQCFP**A*P*T**S*KLLYALTX 215
           ******:****************************************** *********

TSWV-PK-2  M*FKVQKCNSVSRSRFLGF*CDDL*D 238
TSWV-PK-1  M*FKVRKCNSVFRSRFLGF*CDDL*D 238
TSWV-PK-5  M*FKVQKCNSVSRSRFLGF*CDDL*D 238
TSWV-PK-3  M*FKVRKCNSVFRSRFLGF*CDDL*D 238
TSWV-PK-4  M*FKVRKCNSVFRSRFLGF*CDDL*D 238
           *****:***** **************
```

Fig. 1: Amino acid sequence alignment of five Pakistani *Tomato spotted wilt virus* isolates

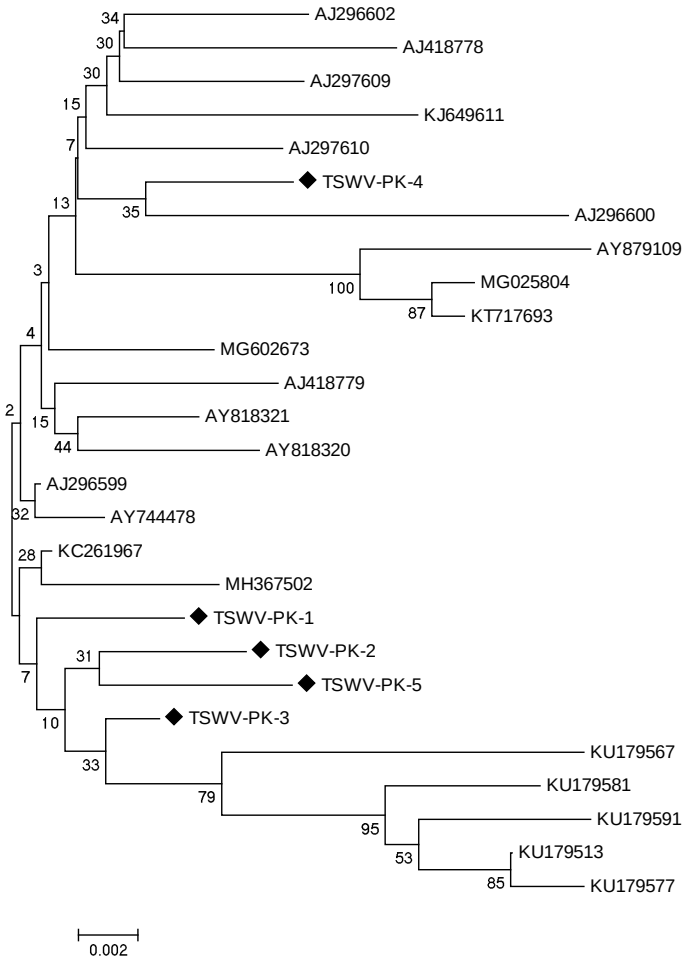

Fig. 2: A 1000 bootstrap consensus tree based on nucleotide sequences of small (S) fragment; partial nucleocapsid (N-gene) of five *Tomato spotted wilt virus* isolates from Pakistan (dotted) and 22 database sequences

Ethics Approval

Not applicable.

References

Adkins S (2000). *Tomato spotted wilt virus*-positive steps towards negative success. *Mol Plant Pathol* 1:151–157

Antignus Y, M Lapidot, N Ganaim, J Cohen, O Lachman, M Pearlsman, BA Raccah (1997). Biological and molecular characterization of *Tomato spotted wilt virus* in Israel. *Phytoparasitica* 25:319–330

Balasubramanian V, R Selvarajan (2014). Genetic diversity and recombination analysis in the coat protein gene of *Banana bract mosaic virus*. *Vir Genes* 48:509–517

Ciuffo M, L Tavella, D Pacifico, V Masenga, M Turina (2008). A member of a new Tospovirus species isolated in Italy from wild buckwheat (*Polygonum convolvulus*). *Arch Virol* 153:2059–2068

Hanssen IM, M Lapidot, BP Thomma (2010). Emerging viral diseases of tomato crops. *Mol Plant Microb Interact* 23:539–548

Ismaeil F, HAJ Kaseem, AA Al-Chaabi, SA Abdulkader, M Alkhalaf (2015). Serological and molecular characterization of Syrian *Tomato spotted wilt virus* isolates. *Phytopathol Mediterr* 54:28–34

Larkin MA, G Blackshields, N Brown, R Chenna, PA McGettigan, H McWilliam, F Valentin, IM Wallace, A Wilm, R Lopez (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948

Lauring AS, R Andino (2010). Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog* 6; Article e1001005

Librado P, J Rozas (2009). DnaSP v. 5. A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452

Moury B, V Simon (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

Parrella G, P Gognalons, K Gebre-Selassie, C Vovlas, G Marchoux (2003). An update of the host range of *Tomato spotted wilt virus*. *J Plant Pathol* 85:227–264

Rice P, I Longdenand, A Bleasby (2000). EMBOSS: The European molecular biology open software suite. *Trends Genet* 16:276–277

Rosenberg MS (2005). Evolutionary distance estimation and fidelity of pair wise sequence alignment. *BMC Bioinform* 6:102-110

Soellick TR, JF Uhrig, GL Bucher, JW Kellmann, PH Schreier (2000). The movement protein NSm of *Tomato spotted wilt tospovirus* (TSWV): RNA binding, interaction with the TSWV N protein, and identification of interacting plant proteins. *Proc Natl Acad Sci USA* 97:2373–2378

Takeda A, K Sugiyama, H Nagano, M Mori, M Kaido, K Mise, S Tsuda, T Okuno (2002). Identification of a novel RNA silencing suppressor, NSs protein of *Tomato spotted wilt virus*. *FEBS Lett* 532:75–79

Tamura K, G Stecher, D Peterson, A Filipski, S Kumar (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729

Tsompana M, J Abad, M Purugganan, J Moyer (2005). The molecular population genetics of the *Tomato spotted wilt virus* (TSWV) genome. *Mol Evol* 14:53–66

Xu C, X Sun, A Taylor, C Jiao, Y Xu, X Cai, X Wang, C Ge, G Pan, Q Wang (2017). Diversity, distribution and evolution of tomato viruses in China uncovered by small RNA sequencing. *J Virol* 91; Article e00173-17