This is the peer reviewed version of the following article

# Fontana A, Albanese G, Manglli A, Tomassoli L, Tiberini A, 2019. phylogenetic analysis based on full genome sequencing of Italian Tomato Spotted Wilt Virus isolates identified in 'Roggianese' sweet pepper and chili pepper, Annals of Applied Biology 2020, Volume 176, Pages 170-179, ISSN: 1744-7348

which has been published in final doi: 10.1111/aab.12566 (https://onlinelibrary.wiley.com/doi/pdf/10.1111/aab.12566)

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website

Annals of Applied Biology

Annals of Applied Biology An international journal of the QQD



# PHYLOGENETIC ANALYSIS BASED ON FULL GENOME SEQUENCING OF ITALIAN TOMATO SPOTTED WILT VIRUS ISOLATES IDENTIFIED IN 'ROGGIANESE' SWEET PEPPER AND CHILI PEPPER

Journal:	Annals of Applied Biology
Manuscript ID	AAB-2019-0240.R2
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Fontana, Anna; Universita degli Studi Mediterranea di Reggio Calabria, Department of AGRARIA; Universita degli Studi Mediterranea di Reggio Calabria, Mediterranea University of Reggio Calabria Albanese, Giuliana; Universita degli Studi Mediterranea di Reggio Calabria, Dipartimento di AGRARIA Manglli, Ariana; Consiglio per la Rricerca in Agricoltura e l'Analisi dell'Economia Agraria , Research Centre of Plant Control and Certification Tomassoli, Laura; Consiglio per la Rricerca in Agricoltura e l'Analisi dell'Economia Agraria , Research Centre of Plant Control and Certification Tomassoli, Laura; Consiglio per la Rricerca in Agricoltura e l'Analisi dell'Economia Agraria , Research Centre of Plant Control and Certification Tiberini, Antonio; Universita degli Studi Mediterranea di Reggio Calabria, AGRARIA
Key Words:	TSWV, Capsicum chinense, Molecular phylogeny, Molecular evolution, Capsicum annuum



# PHYLOGENETIC ANALYSIS BASED ON FULL GENOME SEQUENCING OF ITALIAN TOMATO SPOTTED WILT VIRUS ISOLATES IDENTIFIED IN 'ROGGIANESE' SWEET PEPPER AND CHILI PEPPER

Anna Fontana<sup>1</sup>, Giuliana Albanese<sup>1</sup>, Ariana Manglli<sup>2</sup>, Laura Tomassoli<sup>2</sup>, Antonio Tiberini<sup>1</sup>

<sup>1</sup>Department of AGRARIA, Mediterranean University of Reggio Calabria, Feo de Vito, 89122

7 Reggio Calabria, Italy. <sup>2</sup>Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia

8 Agraria, Research Centre of Plant Control and Certification, Rome, Italy

### 9 ABSTRACT

A study aimed at defining population structure of Italian tomato spotted wilt virus (TSWV) isolates was performed. Full genome sequencing of six TSWV isolates found in two Italian regions (two from Latium: Lazio 17 and Tarquinia; and four from Calabria: PepCal 10, 12, 22 and 24) were assembled. Identity percentages in nucleotide sequence among these TSWV isolates are here provided. The six full length genome sequences were compared with other two Italian isolates (p105 and p202/3WT) already fully sequenced, as well as full TSWV genomes that could be retrieved from GenBank. Phylogenetic analysis, performed in concatenated sequences and for each gene of each genome segment (L, M and S), confirmed the presence of two clades, namely A-like and D-like. In particular, the phylogenetic tree based on segment L grouped all the newly sequenced TSWV isolates in D-like clade. In the M segment phylogenetic tree, all our TSWV isolates shifted in the A-like clade. Isolates separation was not correlated to their geographical origin in phylogenetic study of distinct ORFs encoded by the RNA S segment. In fact, in nucleocapsid-encoding phylogenetic tree, PepCal 10 and 22 grouped in an A-like clade with p105, PepCal 12 and 24 in a D-like clade with p202/3WT, whereas Lazio 17 and Tarquinia in a third well distinct group. NSs tree displayed only PepCal 10 with p105 in A-like clade, whereas PepCal 12, 22, 24 with p202/3WT in D-like subclade; and isolates from Latium grouped a separated clade adjacent to D-like isolates. Additional analysis on putative reassortment events showed that TSWV Calabrian isolates likely originated from a reassortment event in M RNA and other in S RNA with p105 as major parent. Recombination events were detected in isolates from Latium in L and S RNAs with Chinese isolates as putative major parent.

58 31 *Keywords:* TSWV, *Capsicum annuum*, *Capsicum chinense*, Molecular phylogeny,
 60 32 Molecular evolution.

#### INTRODUCTION

Tomato spotted wilt virus (TSWV) belongs to genus Orthotospovirus, (family Tospoviridae, order Bunyavirales) and is present in all countries with temperate, tropical and subtropical climate conditions (Mateus et al., 2012). Even though a disorder attributed to TSWV was observed in 1906 (Sakimura, 1962), the first description of this disease, detected on tomatoes in 1915 in the state of Victoria (Australia) and named as "spotted wilt of tomato", was published by Brittlebank (1919). A specific characterization of a virus as causal agent of the above disease was reported by Samuel et al. (1930) who gave it the name "tomato spotted wilt virus". TSWV was first reported in Europe in the United Kingdom in 1929 in ornamental cherry tomato (Smith, 1932), and thereafter has caused problems repeatedly in vegetable crops, such as tomato, lettuce, pepper, tobacco and in ornamentals especially in southern Europe (Pappu et al., 2009). Symptoms vary depending on the host plant, season and environmental conditions, outdoor-or greenhouse cultivations, and include stunting, necrosis, chlorosis, ringspot or linear patterns affecting leaves, stems and fruit (German et al., 1992; Mumford et al., 1996). This virus is transmitted by several species in the genus Thrips (order Thysanoptera), in a persistent manner, but most efficiently by the western flower thrips, Frankliniella occidentalis (Pergande), whose spreading contributed to the worldwide occurrence of TSWV and tomato spotted wilt disease (Kirk et al., 2003). 

In Italy, TSWV was first detected in ornamentals and vegetables in 1989 (Bellardi and Vicchi, 1990). Since then, TSWV outbreaks were annually reported throughout the country infecting mainly vegetables of the family Solanaceae (Finetti-Sialer et al., 2002). In fact, despite TSWV could affect a wide range of host plants, including 900 plant species (Hanssen et al., 2010), this virus represents the most economically important virus for tomato and pepper in the Mediterranean basin (Turina et al., 2012). Despite the availability of integrated management approaches for containing TSWV disease, including the use of resistant varieties (Jones, 2004; Momol et al., 2004), emergence of TSWV resistance-breaking strains alarmed producers and stakeholders (Roggero et al., 2002; Aramburu et al., 2010; Margaria et al., 2015). 

TSWV genome consists of three single-stranded RNA segments, named L (8.9 kb), M (4.8 kb), and S (2.9 kb) on the basis of their size (Prins et al., 1998). The L RNA is a negative-sense RNA that encodes the RNA-dependent RNA polymerase (RdRp), which plays an essential role in viral replication (de Haan et al., 1991). In contrast, both M and S segments are ambisense RNAs with two open reading frames (ORFs) each, encoding from the viral sense segments and complementary strands. The M RNA encodes the GnGc glycoprotein and the non-structural movement protein (NSm) required for viral cell-to-cell movement (Lian et al.,

#### Annals of Applied Biology

2013). S RNA encodes the non-structural protein (NSs), involved in gene silencing suppression
and required for insect transmission (Kormelink et al., 1991; Margaria et al., 2014a; Sonoda et
al., 2005), and the nucleocapsid (N) protein (Bucher et al., 2003). In addition, both the M RNA
and S RNA possess characteristic A-U rich intergenic regions (IGR) capable of forming stable
hairpin structure (de Haan et al., 1990).

In view of genome complexity and capability of its rearrangement, and in consideration
of emerging concerns about resistance-breaking due to strains able to overcome *Tsw* and *Sw5*resistance genes in pepper and tomato, respectively (Turina et al., 2016), a study aimed at better
evaluating population structure of Italian TSWV isolates was performed.

A previous study conducted on Italian TSWV population, based on NSm gene analysis, highlighted the presence of two distinct phylogenetic lineages associated to two subgroups, denoted TSWV A-like and TSWV D-like (Finetti-Sialer et al., 2002). According to this study, Italian isolates clustered in both clades (A-like and D-like). A more recent study (Margaria et al., 2014b), obtained the full genome sequences of two Italian isolates, retrieved in 1990 and 1999, in Sicily (p202/3WT) and Liguria (p105), respectively. The whole genome comparison of these two isolates confirmed the presence of the two groups in Italy, and in particular, the presence of clade A-like isolates able to overcome plant resistance to TSWV (Roggero et al., 2002; Ciuffo et al., 2005; Crescenzi et al., 2015). 

The present study, in the framework of two projects dealing with surveys of chili pepper (Capsicum spp.) and sweet pepper 'Roggianese' (Capsicum annuum), respectively, reports on the detection, identification, full genome sequencing and analysis of TSWV isolates found in two Italian regions (Latium, central Italy, and Calabria, southern Italy). In these areas, where TSWV is endemic, the virus was specifically reported on chili pepper (*Capsicum* spp.) and sweet pepper 'Roggianese' (Capsicum annuum) (Tomassoli et al., 2014; Fontana et al., 2016). In particular, previous studies showed that TSWV is the most limiting biotic stress of viral aetiology for 'Roggianese' pepper, a local and well appreciated variety in Calabria. In order to analyse the genetic features and variability of TSWV population in these two production areas, four isolates from Calabria ('Roggianese' pepper) and two from Latium (chili pepper), were fully sequenced, by genome walking through RT-PCR. In addition, these isolates were compared with the two Italian isolates (p202/3WT and p105) already fully sequenced, as well as with complete genome sequences that could be retrieved from GenBank. In addition to phylogenetic analysis, studies concerning the molecular diversity and reassortment/recombination events were performed. 

60 34 MATERIALS AND METHODS

#### Plant material and TSWV isolates

During a 2014-2015 survey performed in Calabria and Latium (southern and central Italy) on viruses affecting pepper crops, plants with typical TSWV symptoms (Figure 1) were collected. TSWV identification was performed by DAS-ELISA (LOEWE Biochemica GmbH, Germany). Two and four TSWV isolates from Latium and Calabria, respectively, were chosen for further analysis and full genome sequencing (Table 1). 

# 

# RNA extraction, RT-PCR assay and sequencing

Total RNAs was extracted by grinding 0.5-1.0 gr of leaves in 1 M phosphate buffer pH 7.2 (1:5) in individual plastic bags (Bioreba, Switzerland) and 100 µl of crude extract were used to extract nucleic acids using "Real Total RNA from Tissue and Cell" kit (Durviz, Valencia, Spain) according to the manufacturer's instructions. To obtain the L, M and S segments separately, total RNA extracts were assayed in one-step RT-PCR using primers by Lee et al. (2011). Furthermore, an additional primer pair was designed to amplify the L RNA 3' terminal region (Supplementary Table S1). 

The one-step RT-PCR assays were performed in a mixture containing 2 µl total RNA solution, GoTaq buffer 1X (Promega, Madison, WI, USA) 2.5 mM of each dNTP, 4 µM of each primer, 1.2 U of AMV RT (Promega), 20 U of RNaseOUT (Invitrogen, Carlsbad, CA, USA), 0.75 U of Taq polymerase-Go (Promega) and RNase-free water to the final volume of 25 µL. RT-PCR was performed at 46 °C for 30 min for the cDNA synthesis, followed by 95 °C for 5 min for denaturation and 35 cycles of amplification (94 °C for 1 min, 53 °C for 1 min and 72 °C for 1 min), and a final extension at 72 °C for 10 min. The amplified PCR products were visualized on 1% agarose gel by SYBR Safe - DNA Gel Stain (Thermo Fisher Scientific). To obtain the 3' and 5' un-translated regions (UTR) a rapid amplification of cDNA ends (RACE) was performed using a 5'/3' RACE Kit, 2nd Generation (Roche). PCR products of the expected size were purified using an Amicon®Ultra (Millipor, Billerica MA USA) purification kit, and cloned by the use of pGEM-T Easy Vector (Promega), following manufacturer's instruction. The obtained plasmids were sequenced in both directions using universal primers T7 and SP6 (Bio-Fab Research, Rome, Italy).

#### Sequence comparison and phylogenetic analysis

Alignments and nucleotide comparisons of entire segments (L, M, S), including the intergenic regions (IGRs), of all the isolates sequenced in this study were performed by BLAST and Clustal W software (Thompson et al., 2002) included in MEGA 7 package (Tamura et al., 2007; Kumar et al., 2016).

Page 5 of 23

#### Annals of Applied Biology

Phylogenetic analysis and tree reconstruction were performed for all the concatenated segments (LMS) and for each ORF in each genome segment of all the Italian isolates obtained in this study, including other TSWV full genome sequences retrieved from GenBank. In particular, the two Italian isolates p105 and p202/3WT, previously identified as A-like and D-like isolates, respectively (Margaria et al., 2014b), were included. Aiming to increase the robustness of the above reported results and tree topologies, analysis were conducted using independently three different methods; phylogenetic trees were generated in parallel through the Neighbor-Joining (NJ) and Maximum Likelihood (ML) methods, each with 1000 bootstrap replications and included in MEGA7 software (Tamura et al., 2007), and by SplitsTree 4.14.4 software, using the uncorrected P characters transformation (Huson and Bryant 2006) for the creation of a network visualization. 

## **Recombination Analysis**

Potential genetic exchanges among the sequenced isolates included in this study and other isolates retrieved in GenBank were analysed using Recombination Detection Software v.4.72 (RDP4) (Martin et al., 2010). The analysis was performed using the TSWV full-length concatenated segments of RNA L, RNA M and RNA S. RDP4 implements seven different algorithms to detect recombination events: RDP, Chimaera, BootScan, 3Seq, SiScan, GENECONV and MaxChi. Only recombination events detected by at least five out of seven methods, with a probability value of 0.05 threshold, were considered. RDP4 provides a consensus recombinant score to distinguish the recombinant and parental sequences for each recombination event. 

# **RESULTS**

# 23 Sequence analysis

The whole genome of the six TSWV isolates, four from Calabria (PepCal) and two from Latium (Tarquinia and Lazio 17), on 'Roggianese' pepper and chili pepper, respectively, were obtained and deposited in GenBank with the accession numbers as reported in Supplementary Table S2. Sequence similarity and sequence alignments of full length of L, M, S segments were performed by Blast analysis and Clustal X, respectively. Thus, the complete sequences of the five genes (RdRp, NSm, GnGc, NSs, N) and the intergenic regions (IGRs) in the M and S segments were obtained. Nucleotide full-lengths of all isolates and of each segment/ORF/IGR are displayed in Supplementary Table S2. 

In details, all isolates showed an identical L segment length (8909 nt) with a single ORF
in the antiviral sense of 8640 nt in length (position 8876 – 237) encoding the RdRp (2879-aa).
On the contrary, length variability of both the M and S segments was observed: from 4763 nt

(PepCal 12) to 4772 nt (Lazio 17) and from 2879 nt (PepCal 10) to 2961 nt (PepCal 12) respectively. Anyhow, analysis of each coding ORF showed identical size for the corresponding gene among the isolates. In particular, in the M segment, the ORF in the viral sense (909 nt) encoded for NSm (302 aa), the ORF in the complementary-sense (3408 nt) encoded for GnGc (1135 aa). The S RNA segment encoded for N protein (259 aa) by an ORF in complementary-sense (777 nt) and for NSs (468 aa) by ORF in viral sense (1404 nt). In addition, IGR lengths between NSm and GnGc (IGR<sub>M</sub> on segment M) and between the NSs and N (IGR<sub>S</sub> on segment S) were reported in Supplementary Table S2.

Pairwise nucleotide sequence identities of the six Italian isolates obtained in the present study, determined by ClustalW alignment, ranged from 95.41% to 99.88% for segment L, from 95.74% to 99.29% for segment M and 93.85% - 99.43% for segment S (Supplementary Table S3). In addition, the nucleotide percentage identities of the intergenic regions  $IGR_M$  and  $IGR_S$ , and in each single ORF were calculated and summarised in Supplementary Table S4. In details, IGR<sub>M</sub> and IGR<sub>S</sub> ranged from 82.76% to 98.85% and from 79.74% to 99.81, respectively, showing the highest identity between PepCal 12 and PepCal 24 isolates. Regarding the ORFs, NSm gene showed the highest nucleotide identity (99.98%) between PepCal 22 and PepCal 24 isolates whereas the lowest nucleotide identity was found in the NSs gene between the PepCal 10 and PepCal 24 isolates (92.95%) (Supplementary Table S4). 

# 20 Phylogenetic relationships of TSWV isolates

To determine the phylogenetic relationships among the six TSWV isolates sequenced in this study, including other Italian isolates and full genome sequences retrieved from GenBank, a phylogenetic tree was obtained concatenating sequences of RNA L, RNA M and RNA S (Figure 2). The tree was constructed by MEGA software, using Neighbour-Joining method, including Bootstrap test (1000 reps). The tree topology confirmed the presence of two well distinct clades corresponding to A-like and D-like subgroups (Finetti-Sialer et al., 2002) and specifically including in the A-like clade some TSWV isolates reported to be able to break Sw5 gene resistance (Margaria et al., 2015; Debreczeni et al., 2015). As shown in Figure 2, the Italian isolates from Calabria and Latium grouped separately into two distant clades. In particular, isolates from Latium (Lazio 17 and Tarquinia) clustered in the A-like subgroup along with Chinese sequences in a well-supported subclade A<sub>1</sub>, distinct from the representative P105 isolate subclade A2, whereas isolates from Calabria (PepCal 10, 12, 22 and 24) joined the D-like subgroup forming a specific evolutionary subclade D<sub>2</sub> more close to p202/3WT isolate (subclade  $D_1$ ) than Korean isolates (subclade  $D_3$ ).

#### Annals of Applied Biology

Further, for each segment (L, M, S) and ORF thereof identified, phylogenetic analyses
were performed by Neighbour Joining method including sequences retrieved from GenBank of
full-length genome and available genes of TSWV isolates, from Italy, Mediterranean countries
and worldwide (Figure 3).

Tree topologies obtained by each ORF using Maximum Likelihood analysis (implemented in MEGA software), for the TSWV segments L, M and S (Supplementary Figure S1), confirmed the above reported clear demarcation in three main clades, except for the RdRp tree. In fact, in this phylogenetic tree (Supplementary Figure S1 L) Italian isolates were confirmed to join the D-like group, as reported in RdRp NJ analysis, without the clear demarcation of isolates from Calabria and from Latium. In particular, by both phylogenetic analyses, the tree topology of segment L, including available RdRp sequences, showed an overall division in two main clades: in the first one there were gathered the Italian A-like p105 isolate, others from USA, South Korea, one isolate from Spain (PVR) but none of the isolates here sequenced. The second clade included the Italian reference D-like p202/3WT isolate together with all the Calabrian isolates (PepCal 10, 12, 22 and 24), other isolates from Spain, Australia and South Korea. In addition, the two isolates from Latium (Lazio 17 and Tarquinia) clustered with BR-01 isolate from Brazil (Figure 3, L, Supplementary Figure S1, L). However, a well distinct group, formed close to the D-like clade, was obtained only in NJ analysis). 

Segment M phylogenetic trees, both by NJ and ML methods, constructed for ORFs coding GnGc and NSm proteins, respectively, showed a similar topology consisting of the two main clades according the TSWV subgroups. The D-like clade included the Italian p202/3WT together with TSWV-D, Br01, Australian, Brazilian, South Korean and Spanish isolates; the A-like clade included all the four Calabrian isolates together with p105 and TSWV isolates from the USA and South Korea; isolates from Latium grouped in a distinct subclade with Chinese isolates, close to A-like clade (Figure 3, M, Supplementary Figure S1, M).

The tree topology obtained analysing the S segment in both genes N and NSs, and with both NJ and ML methods, showed relevant differences between the two ORFs. In phylogenetic tree constructed with N gene sequences, it is possible to identify a third main clade other than those identified as A-like and D-like subgroup for the presence of the reference Italian isolates, p105 and p202/3WT, respectively. Calabrian isolates split into two different clades, with PepCal 12 and PepCal 24 joining the Italian isolate p202/3WT and TSWV-D in D-like subgroup, whereas PepCal 10 and PepCal 22 clustered together with the other Italian isolates p105 and BR-01 in the A-like subgroup. The third clade, gathering Lazio 17 and Tarquina, was clearly separated by the other two. The analysis performed using gene NSs showed a tree

topology similar to that observed for the RdRp gene, where two subclades could be identified
for the D-like subgroup. Further, Calabrian isolates grouped in different clusters compared to
the above analysis as PepCal 22 remains together with PepCal 12, PepCal 24 and p202/3WT in
D-like, while PepCal-10 joined p105 and BR-01 in the A-like subgroup. Lastly, the isolates
from Latium confirmed the close genetic relationship with isolates from China in a distinct subclade close to D-like group (Figure 3, S Supplementary Figure S1, S).

The networking analysis performed by SplitsTree software, generally confirmed what obtained using NJ and ML methods as phylogenetic tree reconstruction analysis. In particular, in both LMS concatenated segment and each ORFs included (Supplementary Figure S2 and S3), a clear separation in three main groups could be observed. In LMS and L segment reconstructions was confirmed that Calabrian isolates clustered in D-like clade along with p202/3WT, well distinct from A-like ones. Isolates from Latium (Lazio17 and Tarquinia) joined a distinct group along with Chinese isolates, confirming the analysis performed by NJ method (Figure 3, L). The SplitsTree analysis of the M segment showed a three-way network pattern, confirming the shift of Calabrian isolates with the other Italian isolates p105 (A-like). Lazio17 and Tarquinia still clustered separately. In S segment, again the grouping in three separated clades was confirmed in both N and NSs genes, with isolates from Latium clearly distinct. 

#### *Recombination analysis*

The recombination/reassortment analysis performed by Recombination Detection Program 4.16 (RDP4) software were conducted on the six Italian isolates of this study and all the fully sequenced isolates retrieved from GenBank, using as input file the Clustal W alignment of the concatenation of L, M, S genomic segments as reported in Lian et al. (2013). The analysis results are reported in Table 2, where for each putative event information is shown. The identification of the reassortment or recombination, position, size of segment, and if other isolates occurred to the same event as major/minor parental. Additional information is reported in Supplementary Table S5 including the consensus obtained by each of the six different algorithms (GENECONV, Bootscan, Chimaera, MaxChi, SiScan, 3Seq and RDP).

The analyses identified three putative reassortment events and four putative recombination events. Among the putative reassortment events, 2 out of 3 involved the Italian isolates obtained in this study. In details, in the first putative event PepCal 12, and the other Calabrian isolates, resulted in a reassortment of segment M (positions 8769-13748 nt), confirmed by all algorithms, with p105 (A-like) and K3-CY-CN as major and minor parental, respectively. In addition, confirmed by all algorithms, PepCal 10 and PepCal 22 showed a

second putative reassortment event in segment S (pos. 13801-16868 nt), with p105 as major and PepCal 24 as minor parental. Tarquina and Lazio 17 showed a putative recombination event in segment L (pos. 2085-2234 nt), confirmed by 7/7 algorithms, with Chinese group and Australia sequences as major and minor parental, respectively. Furthermore, Lazio 17 showed a second putative recombinant event in the S segment (15791-16062 nt) again with Chinese group as putative major parent and with South Africa isolate as putative minor parent. This event was confirmed only in 5 out of 7 algorithms. Other two putative recombination events occurred in the Calabrian isolates, one in PepCal 22 in segment S (pos. 14015-14904 nt) (major: PepCal 10; minor: PepCal 12) and one in segment L (pos. 8666-8768 nt) in all the isolates from Calabria (major: p202/3WT; minor: Chinese group). These putative events were confirmed by 7/7 and 4/7 algorithms, respectively. 

#### **DISCUSSION**

TSWV is an important re-emerging plant virus, occurs in a great number of crops, and represents for most of them a major constraint (Lian et al., 2013). TSWV intrinsic features (such as the multipartite genome), allowing full-length genomic segment exchange among isolates, highlight phylogenetic variability among the three RNA segments (Temchev et al., 2011), indicating the occurrence of genomic reassortment. This is known to play an important role in TSWV evolution (Margaria et al., 2015), in particular in view of the emerging of resistance breaking strains, capable to overcome Tsw and Sw5 resistance genes in pepper and tomato, respectively. 

In view of the reassortment role of segment M in resistance breaking (Hoffmann 2001), there is the need of studies regarding TSWV populations and dynamics. Previous studies reported the presence in Italy of isolates belonging to the two strains (Finetti-Sialer et al., 2002; Margaria et al., 2014b) associated to subgroups D and A (Jahn et al., 2000).

In this study, the whole genome sequences of six TSWV isolates detected in pepper plants from two Italian regions (Latium and Calabria, from chili and sweet pepper, respectively) were determined. Comparative analysis of whole-genome sequences indicated a clear demarcation among isolates from Latium (chili pepper) and Calabria (sweet pepper).

Further, these isolates were compared to other TSWV isolates retrieved from GenBank,
from Italy, Mediterranean basin and worldwide. In particular, Italian and Spanish isolates
representatives of either subgroup D (namely D-like) or subgroup A (namely A-like) (Margaria
et al., 2015; Debreczeni et al., 2015), responsible for breaking resistance in tomato (*Sw5*) and
pepper (*Tsw*), respectively, were included. The phylogenetic analyses of each ORF in L, M and

S segments, obtained independently applying different software and algorithms (NJ, ML and
SplitsTreee), confirmed what was reported in Debreczeni et al. (2015), with an overall clear
topology, with two main clades associated to subgroup D and A (Margaria et al., 2015), each
including the Italian reference p105 and p202/3Wt isolates, respectively.

Italian isolates obtained in this study showed a different grouping depending on the
segment analysed. In all tree topology reconstruction methods (NJ, ML and SplitsTree), Latium
isolates confirmed to cluster separately, along with isolates from China, suggesting a different
origin from the other Italian isolates included in this study, and therefore forming a sub-clade,
clearly visualized in SplitsTree LMS tree (Supplementary Figures S2 and S3).

This "Chinese" TSWV flow gene hypothesis could be strengthened by reports of emerging viruses on chili pepper cultivation in Latium as pepper vein yellows virus (PeVYV) and chili veinal mottle virus (ChiVMV) (Tomassoli et al., 2016; Tiberini et al., 2017). In both cases, phylogenetic analysis revealed a Chinese origin, probably due to the fact that tomato and pepper production, breeding and genetic improvement of major companies are based in China. In addition, recent introduction and establishment of vegetable and food commodities by Chinese communities were reported (Tiberini et al., 2017), increasing the risk of introducing new viruses.

As above reported, confirmed independently by three phylogenetic analysis (NJ, ML and SplitsTree), despite the fact that Calabrian isolates grouped with Italian isolate p202/3WT and Pujol1TL3 (Spain) isolates (belonging to D-like group) in concatenated genome LMS, in L and partially in S segment analysis a clear different clustering was observed in M segment. In fact, all the Calabrian isolates joined the A-like group along with p105 and PVR (Spain). This shifted phylogenetic relationship occurred in both M segment ORFs (GnGc and NSm), suggesting a hypothetical reassortment of the segment M in Calabrian isolates (PepCal 10, PepCal 12, PepCal 22, PepCal 24). In addition, a putative reassorment and recombination event could be also predicted for the S segment for PepCal 10 and PepCal 22, respectively; in fact, PepCal 10 clustered in A-like clade in both ORFs phylogenetic trees reconstruction, whereas for PepCal 22 occurred only in N protein ORF tree (Figure 3, S, Supplementary Figures S1 and S3). 

These events, deducted from tree topology, were confirmed by recombination analysis using RDP4 software using as input file the full-length concatenated L, M and S segments. The accordance among three different tree reconstruction methods and recombinant analysis confirmed that L, M and S segments of Italian isolates obtained in this study could have different origins, implying the occurrence of frequent reassortment during virus evolution. In

#### Annals of Applied Biology

particular, Latium isolates could have a similar origin to the Chinese isolates, thereof origin
 was demonstrated to be derived from American and secondary European isolates (Turina et al.,
 2012); on the contrary, the Calabrian population could have passed from a bottleneck of other
 indigenous (different natural hosts) or foreign populations.

Moreover, these data suggest genetic exchanges between Asiatic and European TSWV populations. In fact, the recombination analysis showed additional few recombination events for both isolates from Calabria and Latium. In literature, it is reported how almost all recombination events were mainly detected in the 5' half of the sense RNAs. As previously reported by Zhang et al. (2016) the reason for this preference is not clear. One possibility is that RdRp firstly bound RNA template, synthesized complementary strands and then had chance to facilitate recombination among different complementary strands (Lian et al., 2013). In this case, the reassortment/recombination occurred in M and S segments, this is particularly relevant in view of emerging new resistance breaking isolates. Segment M resulted highly involved in Sw5 resistance overcome in tomato along with the NSs gene of RNA S reported as silencing suppressor (Pappu et al., 2009; Webster et al., 2015; Turina et al., 2016) and virulence determinants in the interaction with *Tsw* resistance gene in pepper (Zhang et al., 2016). This is particularly relevant for PepCal 10, where a reassortment occurred not only in segment M but also in segment S, highlighted by its clustering in A-like clade, along with isolates reported to be able to break resistance in pepper (PVR from Spain). 

### 21 CONCLUSION

These results highlight a constant evolving situation in TSWV population dynamics, especially in Italy, where it was possible to observe the formation of a new, formerly undetected, distinct sub-clade composed mainly by isolates from Latium, along with Chinese isolates. Genetic reassortment was confirmed to be fundamental in TSWV evolution and to overcome resistance genes in tomato and pepper (Turina et al., 2016). More studies should be done to associate this selective pressure to emergence of resistance breaking, focusing mainly on PepCal 10 isolate, according to its clustering in M and S segments, along with PVR isolate, reported to have a key role in resistance breaking and silencing suppression in pepper. TSWV full-genome sequence analysis performed in this study adds useful information to better understand the evolution of new virus isolates and provide a new update of the Italian and Mediterranean TSWV population evolution.

 Acknowledgements

This work was carried out in the frame of a research project: "Tutela e valorizzazione delle produzioni del peperone autoctono Roggianese" supported by Regione Calabria, Italy (PSR Calabria 2007-2013, ASSE 2 "Miglioramento dell'ambiente e dello spazio rurale" Misura 214, Azione 6 "Progetti comprensoriali per la salvaguardia del patrimonio genetico regionale"). COMPLIANCE WITH ETHICAL STANDARDS Ethical statement. This research did not involve any animal and/or human participant. The authors declare that they have no conflict of interest. **BIBLIOGRAPHY** Aramburu J., Galipienso L., Soler S. and Lopez C. (2010) Characterization of Tomato spotted wilt virus isolates that overcome the Sw-5 resistance gene in tomato and fitness assays. Phytopathol. Mediterr. 49:342–351 Bellardi M.G. and Vicchi V. (1990) TSWV una nuova insidia per la produzione agricola italiana. Informatore fitopatologico 60(9): 17-24. Brittlebank, C.C., 1919 Tomato diseases. J. Agr. Victoria., 17: 213-23. Bucher E., Sijen T., De Haan P., Goldbach R., Prins M. (2003) Negative-strand tospoviruses and *tenuiviruses* carry a gene for a suppressor of gene silencing at analogous genomic positions. J Virology 77: 1329-1336. Crescenzi A., Fanigliulo A., Viggiani G. (2015) Resistance breaking tomato spotted wilt virus isolates on resistant tomato cultivars in Italy. Acta horticulturae, 1069: 95-98. Ciuffo M., Finetti-Sialer M. M., Gallitelli D., and Turina M. (2005) First report in Italy of a resistance-breaking strain of Tomato spotted wilt virus infecting tomato cultivars carrying the Sw5 resistance gene. Plant Pathol. 54:564. Debreczeni D.E., Lopez C., Aramburu J., Daros J.A., Soler S. (2015) Complete sequence of three different biotypes of tomato spotted wilt virus (wild type, tomato Sw-5 resistance-breaking and pepper Tsw resistance-breaking) from Spain. Arch. Virol. 160:2117-23. de Haan P., Wagemakers L., Peters D., Goldbach R. (1990) The S RNA segment of Tomato spotted wilt virus has an ambisense character. J Gen Virol.; 71:1001–7. 

1	de Haan P., Kormelink R., Oliveira R., Poelwijk F., Peters D., Goldbach R. (1991) Tomato
2	spotted wilt virus L RNA encodes a putative RNA polymerase. J Gen Virol.;72:2207–16.
3	Finetti-Sialer M.M., Lanace V., Padula M., Vovlas C., Gallitelli D., (2002) Occurrence of
4	two distinct Tomato spotted wild virus (TSWV) subgroups in Southern Italy. J. Plant Pathol.
5	<b>84</b> , 145–152.
6	Fontana A., Lupini L., Tocci C., Zicca F., Garritano D., Tomassoli L., Sunseri F., Albanese
7	G., Monti M. (2016) Il peperone 'Roggianese': caratterizzazione e indagine sui virus.
8	Colture Protette, 12: 52-56.
9	German, T.L., D.E. Ullman and J.W. Moyer (1992) Tospoviruses: diagnosis, molecular
10	biology, phylogeny and vector relationships. Annu. Rev. Phytopathol., 30: 315-348.
11	Hanssen I.M., Lapidot M., Thomma B.P. (2010) Emerging viral diseases of tomato crops.
12	Mol Plant Microbe Interact 23:539–548.
13	Hoffmann, K., Qiu, W. P., and Moyer, J. W. (2001). Overcoming host- and pathogen
14	mediated resistance in tomato and tobacco maps to the M RNA of Tomato spotted wilt
15	virus. Mol. Plant Microbe Interact. 14:242–249.
16	Huson D. H and Bryant D. (2006) Application of Phylogenetic Networks in Evolutionary
17	Studies. Mol. Biol. Evol., 23(2):254-267, 2006
18	Jahn M., Paran I., Hoffmann K., Radwanski E.R., Livingstone K.D. (2000) Genetic
19	mapping of the Tsw locus for resistance to the Tospovirus Tomato spotted wilt virus in
20	Capsicum spp. and its relationship to the Sw-5 gene for resistance to the same pathogen in
21	tomato. Mol. Plant-Microbe Interact. 13:673–82.
22	Jones R.A.C. (2004) Using epidemiological information to develop effective integrated virus
23	disease management strategies. Virus Res 100:5-30
24	Kirk W., Terry L. (2003) The spread of the western flower thrips Frankliniella occidentalis
25	(Pergande). Agric For Entomol. 5: 301–310.
26	Kormelink, R., Kitajima, E. W., De Haan, P., Zuidema, D., Peters, D., and Goldbach, R.
27	(1991) The nonstructural protein (NSs) encoded by the ambisense S RNA segment of tomato
28	spotted wilt virus is associated with fibrous structures in infected plant cells. Virology 181:
29	459–468.
	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29

23	
4	
5 6	
7	
8 9	
10 11	ļ
12	
13 14	
15 16	
17	:
18 19	
20 21	1
22	1
23 24	1
25 26	1
27	1
28 29	14
30 31	
32	1.
33 34	1
35 36	1
37	1
39	1
40 41	2
42 43	
43 44	2
45 46	2
47 48	2
49	2
50 51	-
52 53	2
54	2
55 56	2
57 58	2
59	
60	

Kormelink R, Haan P, Meurs C, Peters D, Goldbach R. (1992) The nucleotide sequence of
 the M RNA segment of *Tomato spotted wilt virus*, a bunyavirus with two ambisense RNA
 segments. *J Gen Virol.* 73:2795–804.

Kumar S., Stecher G., and Tamura K. (2016) MEGA7: Molecular Evolutionary Genetics
Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870-1874.

Lee J., Cho W., Kim M., Kwak H., Choi H. (2011) Complete genome sequences of three
Tomato spotted wilt virus isolates from tomato and pepper plants in Korea and their
phylogenetic relationship to other TSWV isolates. *Arch Virol* 156: 725–728.

9 Lian S., Lee J.S., Cho W.K., Yu J., Kim M.K. (2013) Phylogenetic and Recombination
10 analysis of Tomato Spotted Wilt Virus. *PLoS ONE* 8(5): e63380. doi:
11 10.1371/journal.pone.0063380.

Margaria P., Bosco L., Vallino M., Ciuffo M., Mautino G.C., Tavella L., Turina M. (2014a)
 The NSs protein of tomato spotted wilt virus is required for persistent infection and
 transmission by Frankliniella occidentalis. J. Virol. 88, 5788–5802.

Margaria P., Miozzi L., Ciuffo M., Pappu H., Turina M., (2014b) The first complete genome sequences of two distinct European tomato spotted wilt virus isolates. *Arch. Virol.*, http://dx.doi.org/10.1007/s00705-014-2256-3

Margaria P., Ciuffo M., Rosa C., Turina M. (2015). Evidence of a *Tomato spotted wilt virus* resistance breaking strain originated through natural reassortment between two
 evolutionary-distinct isolates. *Virus Res.* 196:157–61.

Martin D.P., Lemey P., Lott M., Moulton V., Posada D., Lefeuvre P. (2010) RDP3: A
flexible and fast computer program for analyzing recombination. *Bioinformatics* 26:2462–
3.

Mateus C., Pequito A., Teixeira S., Queiro's R., Godinho M.C. (2012) Development of a
 *Tomato spotted wilt virus* (TSWV) risk evaluation methodology for a processing tomato
 region. Span J Agric Res 10: 191–197.

Momol M.T., Olson S.M., Funderburk J.E., Stavisky J., Marois J.J. (2004) Integrated
management of tomato spotted wilt on field-grown tomatoes. *Plant Dis* 88:882–890

1 ว		
2 3	1	Mumford R.A., Barker I., Wood K.R. (1996) The biology of the tospoviruses. Ann Appl
4 5 6	2	<i>Biol.</i> ; <b>128</b> :159–83.
7 8	3	Pappu H.R., Jones R.A.C., Jain R.K. (2009) Global status of tospovirus epidemics in diverse
9	4	cropping systems: Successes achieved and challenges ahead. Virus Research 141: 219-236.
10 11 12	5	doi:10.1016/j.virusres.2009.01.009.
12	6	Prins M Goldbach R (1998) The emerging problem of <i>tospovirus</i> infection and non-
14 15	7	conventional methods of control. <i>Trends Microbiol</i> <b>6</b> : 31–35.
16 17	0	
18 19	8	Roggero P., Masenga V., Tavella L., (2002) Field isolates of Tomato spotted wilt virus
20	9	overcoming resistance in pepper and their spread to other hosts in Italy. <i>Plant Disease</i> 86,
21 22	10	950-954.
23 24	11	Sakimura K. (1962). The present status of thrips borne viruses. In: Biological Transmission
25 26	12	of Disease Agents. K. Maramorosch (ed.). Academic Press, New York, USA, 33-40.
27 28 20	13	Samuel G., Bald J.G. and Pitman H.A. (1930) Spotted wilt of tomatoes. J. Austral. Science
29 30 31	14	Ind. Res., 44.
32 33	15	Smith K.M. (1932) Studies on plant virus diseases. XI. Further experiments with a ringspot
34 35	16	virus: its identification with spotted wilt of tomato. Ann. Appl. Biol. 19: 305-320.
36 37	17	Sonoda S., Nishiguchi M. and Tsumuki H. (2005) Evaluation of Virus Resistance Conferred
38 39	18	by the NSs Gene Sequences from Tomato spotted wilt virus in Transgenic Plants Breeding
40 41	19	<i>Science</i> <b>55</b> : 27–33
42 43	20	Tamura K., Dudley J., Nei M. & Kumar S. (2007) MEGA4: Molecular Evolutionary
44 45	21	Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution
46 47	22	<b>24</b> :1596-1599.
48 49	23	Tentchev D., Verdin E., Marchal C., Jacquet M., Aguilar J., (2011) Evolution and structure
50 51	24	of Tomato spotted wilt virus populations: evidence of extensive reassortment and insights
52 53	25	into emergence processes. J Gen Virol 92: 961–973.
54 55	26	Thompson J., Gibson T., Higgins D. (2002) Multiple sequence alignment using ClustalW
56 57 58 59 60	27	and ClustalX. Curr Protoc Bioinformatics: 2.3. 1–2.3. 22.

2
2
3
4
5
6
7
/
8
9
10
11
10
12
13
14
15
16
10
17
18
19
20
20
21
22
23
24
24
25
26
27
28
20
29
30
31
32
22
22
34
35
36
27
27
38
39
40
41
40
42
43
44
45
16
40
47
48
49
50
50
21
52
53
54
55
55
56
57
58
50

Tiberini A., Manglli A., Ahmad A., Cassia U., Tomassoli L. (2017) First report and molecular
 identification of Chili veinal mottle virus in Italy. *Journal of Plant Pathology* 99(2): 523-526. DOI:
 10.4454/jpp.v99i2.3891

Tomassoli L., Vitale S., Tiberini A. e Barba M. (2014) Il peperoncino: presenza di malattie
negli areali di coltivazione nel Centro Italia. *Colture Protette*, 4: 48-54.

Tomassoli L, Manglli A, Ahmad A, Tiberini A, Barba M. (2016) First report of Pepper
vein yellows virus infecting chili pepper (Capsicum spp.) in Italy. *New Disease Reports* 33, 22. http://dx.doi.org/10.5197/j.2044-0588.2016.033.022

9 Turina M., Tavella L., Ciuffo M. (2012) Tospoviruses in the Mediterranean area. In:
10 Loebenstein G, Lecoq H (eds) Viruses and virus diseases of vegetables in the Mediterranean
11 Basin, Advances in Virus Research. Elsevier Academic Press Inc, San Diego, pp 403–437.

Turina M., Kormelink R., and Resende R.O. (2016) Resistance to tospoviruses in vegetable
crops: epidemiological and molecular aspects. *Annu. Rev. Phytopathol.* 54:347-71.

Webster C.G., Frantz G., Reitz S.R., Funderburk J.E., Mellinger H.C. (2015) Emergence of
 *Groundnut ringspot virus* and *Tomato chlorotic spot virus* in vegetables in Florida and the
 south eastern United States. *Phytopathology* 105:388–98

Zhang Z., Wang D., Yu C., Wang Z., Dong J., Shi K., Yuan X (2016) Identification of three
 new isolates of *Tomato spotted wilt virus* from different hosts in China: molecular diversity,
 phylogenetic and recombination analyses. *Virology Journal* 13:8 DOI 10.1186/s12985-015 0457-3.

20

60

 Figure 1. TSWV symptoms on 'Roggianese' pepper and 'Habanero Francisca' chili pepper on left and right panel,
 respectively.

Figure 2. Phylogenetic tree of full-length TSWV isolates constructed by MEGA software (Tamura et al., 2007), using Neighbour-Joining method, with Bootstrap test (1000 reps). The whole-nucleotide sequences were obtained concatenating RNA L, RNA M, and RNA S sequences of the six TSWV isolates (two from Latium: Lazio 17 and Tarquinia; and four from Calabria: PepCal 10, 12, 22 and 24) sequenced in this study and other full genome sequences retrieved from GenBank. PepCal 10 (L: MH763621; M: MH756624; S: MG989673); PepCal 12 (L: MK348941; M: MH756625; S: MG989674); PepCal 22 (L: MH763622; M: MH756626; S: MG989675); PepCal 24 (L: MH763623; M: MH756627; S: MG989676); Tarquinia (L: MK348942; M: MG983522; S: MG983521); Latium 17 (L: MK348943; M: MG983520; S: MG983519).

**Figure 3.** Phylogenetic trees of TSWV ORFs (RdRp, GcGn, NSm, N and NSs) located in RNA L, RNA M, and

12 RNA S segments constructed by MEGA software (Tamura *et al.*, 2007), using Neighbour-Joining method, with

13 Bootstrap test (1000 reps). The sequences of the six TSWV isolates (two from Latium: Lazio 17 and Tarquinia;

14 and four from Calabria: PepCal 10, 12, 22 and 24) were assembled in this study and the others retrieved from

Peer Peurez

15 GenBank.











# Annals of Applied Biology

Table 1- Italian region and type of pepper (Capsicum annuum L. and Capsicum chinense) where TSWV isolates were detected.

	Table 1- Italian	region and type of pepper (Capsicum annuun	n L. and <i>Capsicum</i>	chinense) where TSWV isolates were detec
N°	Name of Isolate	Pepper biotype or variety	Species	Geographic origin
1	Lazio 17	Chilli pepper (Habanero Red)	C. chinense	Latium-Roma (RM)
2	Tarquinia	Chilli pepper (Habanero Francisca)	C. chinense	Latium- Viterbo (VT)
3	PepCal 10	Elongated pepper 'Roggianese'	C. annuum L.	Calabria – Roggiano Gravina (CS)
4	PepCal 12	Rounded pepper 'Roggianese'	C. annuum L.	Calabria – Roggiano Gravina (CS)
5	PepCal 22	Elongated pepper 'Roggianese'	C. annuum L.	Calabria – Roggiano Gravina (CS)
6	PepCal 24	Elongated pepper 'Roggianese'	C. annuum L.	Calabria – Roggiano Gravina (CS)
,				1
ł				

- 45 46

#### Annals of Applied Biology

**Table 2**. Recombination/reassortment putative events identified by Recombination Detection Program 4.16 (RDP4) software in this study. RDP4 implements seven different algorithms to detect recombination events: RDP, Chimaera, BootScan, 3Seq, SiScan, GENECONV and MaxChi. For each putative event is reported type of event (reassortment/recombination), other isolates involved, parental (major and minor), position and segment involved. In addition, the last column is reported how many algorithms confirmed independently the putative event.

Event	Isolate	Reass*/Rec+	Other Isolates	Parental	Position	Segment	Coverage
1	PepCal 12	*Reass	PepCal 10	Major: p105	8769-13748	М	7/7
			PepCal 22	Minor: K3-CY-CN			
			PepCal 24				
2	PepCal 10	*Reass	PepCal 22	Major: p105	13801-16868	S	7/7
				Minor: PepCal 24			
3	p202/3WT	*Reass	TSWV_12	Major: TSWV_6	8972-14010	М	5/7
			TSWV_17	Minor: PA (p105-like)			
			SouthAfrica				
			Australia				
5	Tarquinia	+Rec	Latium 17	Major: Chinese group	2085-2234	L	7/7
				Minor: Australia			
6	PepCal 22	+Rec	-	Major: PepCal 10	14015-14904	S	7/7
				Minor: PepCal 12			
12	PepCal 24	+Rec	PepCal 10	Major: p202/3WT	8666-8768	L	4/7
			PepCal 12	Minor: Tarquina			
					2		

Page 23 of 23

Annals of Applied Biology

			PepCal 22	(Chinese group)			
13	Latium 17	+Rec		Major: Chinese group	15791-16062	S	5/7
				Minor: SouthAfrica			
					3		