Plant Pathology (2010) **59**, 785–793 Doi: 10.1111/j.1365-3059.2010.02280.x

Molecular variability of Spanish and Hungarian isolates of *Tomato torrado virus*

A. Alfaro-Fernández[a]*, M. C. Cebrián[a], J. A. Herrera-Vásquez[a], M. C. Córdoba-Sellés[a], J. A. Sánchez-Navarro[b] and C. Jordá[a]

[a]*Instituto Agroforestal Mediterráneo (IAM), Universidad Politécnica de Valencia (UPV), Camino de Vera s/n, 46022 Valencia; and* [b]*Instituto de Biología Molecular y Celular de Plantas (IBMCP), Universidad Politécnica de Valencia UPV-CSIC, Calle Ingeniero Fausto Elio s/n, 46022 Valencia, Spain*

The population structure and genetic variation of *Tomato torrado virus* (ToTV) were estimated from 19 Spanish isolates collected from 2001 to 2009 in different tomato-production areas by analyses of the partial nucleotide sequences of five regions of the virus genome: the protease cofactor (Pro-Co) and the RNA-dependent RNA polymerase (RdRp) in RNA1, and the movement protein (MP) and two subunits of the coat protein (CP; viz. Vp35 and Vp23) in RNA2. Three Hungarian isolates of the virus were also included in the analyses. All the ToTV isolates clustered together in the phylogenetic analysis of the nucleotide sequences of the different regions. However, some genetic diversity was observed in the case of the two CP subunits among the Gran Canaria isolates and the remaining ToTV-isolates analysed, which grouped together. A high similarity was observed among all the isolates and the two published ToTV isolates: the ToTV type isolate (PRI-ToTV0301) and the Polish isolate Wal03. The most variable encoding regions studied were those on RNA2. In general, no correlation was found between genetic diversity and collection date. Studying the genetic distances between pairs of sequences, the ratio between nonsynonymous (amino-acid-replacing) and synonymous (silent mutational) substitutions was low, indicating a strong negative selective pressure in the studied regions. Nine negatively selected sites (distributed in Pro-Co, MP, Vp23 and Vp35) and just one positively selected one (in Pro-Co) were found for all the genome regions studied.

Keywords: genetic variation, phylogenetic analysis, RNA sequence, *Solanum lycopersicum,* weeds

Introduction

Tomato torrado virus (ToTV) is a newly characterized picorna-like virus, and the type member of a new genus, *Torradovirus* (Verbeek *et al.*, 2007a), which also includes two other recently characterized virus isolates of tomato (*Solanum lycopersicum*) namely Tomato apex necrosis virus (ToANV) and *Tomato marchitez virus* (ToMarV) (Turina *et al.*, 2007; Verbeek *et al.*, 2007b; Sanfaçon *et al.*, 2009). Virions of ToTV consist of isometric particles with a diameter of 28 nm and their genome is composed of two single-stranded positive-sense RNA molecules. RNA1 has approximately 7800 nt and contains one open reading frame (ORF) encoding the characteristic functional domains of the protease cofactor (Pro-Co), helicase, protease and the RNA-dependent RNA polymerase (RdRp). RNA2 has approximately 5400 nt and contains two ORFs. ORF2, which partly overlaps ORF1, encodes three subunits of the coat protein (CP) gene, named according to their approximate molecular weights (Vp35, Vp26 and Vp23), and a putative movement protein (MP), since a motif LRVPML, similar to the characteristic region of MP found among other viruses, was determined. ORF1 showed no homology to other viral proteins and the analysis of functional sites did not identify any specific domains (Verbeek *et al.*, 2007a; Budziszewska *et al.*, 2008). This genome structure resembles those of the members of the genera *Sequivirus*, *Waikavirus, Sadwavirus* and *Cheravirus*. However, phylogenetic analysis revealed important differences between ToTV and the viruses of these genera (Verbeek *et al.*, 2007a). Recently, a new family, Secoviridae, which would include all those genera as well as *Comovirus*, *Nepovirus*, *Fabavirus* and *Torradovirus*, has been established (Sanfaçon *et al.*, 2009).

ToTV causes a tomato disease which, in Spain, is locally named 'torrado' because of the necrotic burn-like symptoms on leaves and fruits that affected plants show. The cytopathology of ToTV infection alone and with other viruses is described by Alfaro-Fernández *et al.* (2010). Apart from mainland Spain (Verbeek *et al.*, 2007a), the virus was reported on tomatoes in Poland (Pospieszny *et al.*, 2007), Australia (IPPC, 2008), Panama (Herrera-Vásquez *et al.*, 2009), Hungary

*E-mail: analfer1@doctor.upv.es

Published online 14 April 2010

(Alfaro-Fernández *et al.*, 2009) and France (Verdin *et al.*, 2009), and was recently included in the EPPO alert list (EPPO, 2009). This virus is efficiently transmitted by the whitefly species *Trialeurodes vaporariorum* (Pospieszny *et al.*, 2007) and *Bemisia tabaci* (Amari *et al.*, 2008), and naturally infects several weed species associated with tomato crops (Alfaro-Fernández *et al.*, 2008).

Comparative sequence analyses of the complete genome of two different isolates: the ToTV type isolate PRI-ToTV0301 (Acc. Nos. DQ388879 and DQ388880), collected from the Murcia Region (Spain) in 2003, and the Polish isolate named Wal03 (Acc. Nos. EU563948 and EU563947), collected in the Wielkopolska region (Poland) in 2003, revealed a high homology (98–100% amino acid identity), although few nucleotide and amino acid substitutions were observed in the coding regions (Budziszewska *et al.*, 2008). Other Polish isolates (Kra and Ros) collected in 2007 in another region of the country, also revealed high amino acid identity of three subunits of the CP to each other, as well as in relation to Wal03 and PRI-ToTV0301 (Pospieszny *et al.*, 2010). The comparison of some coding regions of other isolates, mainly fragments containing the MP motif or/and the CP subunit Vp23, also revealed a high level of nucleotide identity (98–99%) with the ToTV type isolate PRI-ToTV0301; these isolates included ToTV-Can (Acc. No. EF436286) isolated from the Canary Islands (Spain) in 2006 (Alfaro-Fernández *et al.*, 2007), ToTV-CE (Acc. Nos. EU476181 and EU476182) from tomato plants in a greenhouse of the Murcia Region (Spain) (Amari *et al.*, 2008), ToTV-Pan1 (Acc. Nos. FJ357161 and EU934037) collected from the Coclé region (Panama) in 2008 (Herrera-Vásquez *et al.*, 2009), ToTV-W1 and ToTV-W2 (Acc. Nos. EU090252 and EU090253, respectively) collected from two weed species from the Murcia Region and the Canary Islands (Spain) in 2007, and the Hungarian isolates ToTV-H1 (EU835496 and FJ616996), ToTV-H2 (FJ616995 and FJ616997) and ToTV-H3 (FJ616994 and FJ616998) collected in Szeged, Öcsöd and Csongrád (Hungary), respectively, in 2007 and 2008 (Alfaro-Fernández *et al.*, 2009).

The aim of the present work was to evaluate the molecular variability of five coding regions of the ToTV genome and the phylogenetic relationships among the analysed Spanish ToTV isolates collected over a period of 9 years (2001–2009) from diverse geographic origins, three Hungarian ToTV isolates and isolates deposited in the GenBank database.

Materials and methods

Virus isolates, RNA extraction and dot-blot hybridization analysis

Leaf samples of plants with symptoms were collected from different geographical regions of Spain and Hungary in the period 2001–2009. The characteristics of the samples are detailed in Table 1. The Spanish isolates collected from tomato were codified by a three-letter code indicative of geographic origin (ALC = Alicante, ALM = Almeria, BCN = Barcelona, GNC = Gran Canaria, MUR = Murcia, TEN = Tenerife), followed by the year of collection. Total RNA extraction was performed with 0·1 g fresh or dried leaf tissue from the infected plants using the silica-capture protocol (MacKenzie *et al.*, 1997). One microlitre of the extracted RNA was first denatured with formaldehyde and then directly applied to a nylon membrane. The analysis of total nucleic acids by non-isotopic dot-blot hybridization was performed as described previously by Sánchez-Navarro *et al.* (1998) using a dig-RNA probe complementary to a fragment of the polyprotein of RNA2 of ToTV which contained the MP conserved motif to verify the presence of the virus.

RT-PCR and sequencing

Five different regions of the virus genome were studied: the partial Pro-Co and RdRp regions in RNA1, and the partial MP and two subunits of the CP (Vp35 and Vp23) in RNA2. Previously described primers were used to amplify these regions, plus three ToTV-specific primers (Vp35-D, CP-R and ToTCoF-R) which were designed using the OLIGO program version 4.0 (National Bioscience Inc.) on the basis of the complete ToTV genome sequences deposited in the GenBank database [ToTV type isolate PRI-ToTV0301 RNA1 (Acc. No. DQ388879) and RNA2 (Acc. No. DQ388880); and Polish isolate Wal03 RNA1 (Acc. No. EU563948) and RNA2 (Acc. No. EU563947)]. Primers and their characteristics are given in Table 2.

RT-PCR reactions of particular fragments of the ToTV genome were performed using the SuperScript III one-step RT-PCR system with the Platinum *Taq* DNA polymerase kit (Invitrogen Life Technologies) with the pair of specific primers (Table 2). The RT-PCR programme consisted of a reverse transcription step at 50°C for 30 min followed by 2 min at 94°C and 40 cycles of 94°C for 15 s, annealing temperature appropriate for specific primers for 30 s and 68°C for 1 min, plus a final incubation at 68°C for 10 min to finish the incomplete PCR fragments. The amplified PCR products were analysed on 1·2% agarose gel in 1× TAE buffer, stained with ethidium bromide and visualized under UV light. The amplified fragments were purified using the High Pure PCR Product Purification Kit (Roche Diagnostics), directly sequenced and deposited in the GenBank database (their accession numbers are provided in Table 1).

Sequence analysis

The similarity/identity of the obtained nucleotide and deduced amino acid sequences was calculated with Matrix Global Alignment Tool software MATGAT version 2.01 (Campanella *et al.*, 2003).

For each genomic region, the model which best described the sequence data was inferred by a maximum likehood approach using the FINDMODEL server (http://

Table 1 Characteristics of the *Tomato torrado virus* (ToTV) isolates used in the assay

				GenBank Acc. No.[b]				
Isolate code[a]	Collection date	Original host	Area of origin	Pro-Co	RdRp	MP	Vp23	Vp35
MUR-01	2001	Tomato	Murcia, Spain	GQ397365	GQ397383[(v)]	GQ397433[(n)]	GQ397399[(s)]	GQ397415
MUR-02	2002	Tomato	Murcia, Spain	GQ397366	GQ397384	[(n)]	GQ397400	GQ397416
TEN-03	2003	Tomato	Tenerife, Spain	GQ397368	GQ397385	GQ397434	GQ397401	GQ397417[(q)]
MUR-03	2003	Tomato	Murcia, Spain	GQ397367[(z)]	[(v)]	GQ397435	[(r)]	GQ397418
W3-04	2004	*Solanum nigrum*	Murcia, Spain	GQ397374	GQ397386[(u)]	GQ397436	GQ397402[(r)]	GQ397419
ALM-04	2004	Tomato	Almería, Spain	GQ397381[(w)]	GQ397387	GQ397437	GQ397403	GQ397420
GNC-04	2004	Tomato	Gran Canaria, Spain	GQ397369[(y)]	GQ397388	GQ397438	GQ397404	GQ397421
MUR-05	2005	Tomato	Murcia, Spain	[(z)]	[(v)]	GQ397439	[(r)]	GQ397422[(p)]
GNC-05	2005	Tomato	Gran Canaria, Spain	GQ397380	GQ397389	GQ397440	GQ397405	GQ397423
GNC-06 (ToTV-CAN)	2006	Tomato	Gran Canaria, Spain	[(y)]	GQ397390	EF436286[c]	GQ397406	GQ397424
MUR-06	2006	Tomato	Murcia, Spain	GQ397377	GQ397391	GQ397441	GQ397407	GQ397425
ALC-07	2007	Tomato	Alicante, Spain	GQ397372	GQ397392	GQ397442	GQ397408	GQ397426[(o)]
W4-07	2007	*Solanum nigrum*	Tenerife, Spain	GQ397371	GQ397393	GQ397445	GQ397409	GQ397427
TEN-07	2007	Tomato	Tenerife, Spain	GQ397378	GQ397394	GQ397444	GQ397410	[(q)]
MUR-07	2007	Tomato	Murcia, Spain	[(w)]	GQ397395	GQ397443	[(s)]	[(o)]
ToTV-H1	2007	Tomato	Szeged, Hungary	GQ397376	[(t)]	EU835496[d]	FJ616996[d]	GQ397430
GNC-08	2008	Tomato	Gran Canaria, Spain	GQ397370	GQ397396	GQ397446	GQ397411	GQ397428
MUR-08	2008	Tomato	Murcia, Spain	GQ397379	[(v)]	GQ397447	GQ397412	[(o)]
BCN-08	2008	Tomato	Barcelona, Spain	GQ397375	[(u)]	GQ397448	GQ397413	GQ397429
ToTV-H2	2008	Tomato	Öcsöd, Hungary	GQ397373[(x)]	GQ397398[(t)]	FJ616995[d]	FJ616997[d]	GQ397431
ToTV-H3	2008	Tomato	Csongrád, Hungary	[(x)]	[(t)]	FJ616994[d]	FJ616998[d]	GQ397432
MUR-09	2009	Tomato	Murcia, Spain	GQ397382	GQ397397	GQ397449	GQ397414	[(p)]

[a]Spanish isolates collected from tomato were codified by three letters indicating geographic origin (ALC = Alicante, ALM = Almeria, BCN = Barcelona, GNC = Gran Canaria, MUR = Murcia, TEN = Tenerife), followed by the year of collection. Isolates from weeds have a 'W' prefix and Hungarian isolates have an 'H' suffix.
[b]Sequences with the same letter (in parentheses) are 100% nt identical.
[c]Sequence published in GenBank and reported by Alfaro-Fernández *et al.* (2007).
[d]Sequences published in GenBank and reported by Alfaro-Fernández *et al.* (2009).

Table 2 Sequences of the *Tomato torrado virus* (ToTV)-specific primers used in the assay

Genome region	Primer name	Nucleotide sequence (5′-3′)	Location in the sequence (nt)[a]	Melting temperature (Tm) (°C)	Expected fragment (bp)	Reference[b]
Pro-Co (RNA1)	ToTGSP2-D[c,d]	GAAGCCTTCGGTTACAGATGCTG	277–299	52	847	Budziszewska *et al.*, 2008
	ToTCoF-R[e]	AGCCTGCTCCCTTAGATGTTGG	1102–1124	52		This study
RdRp (RNA1)	TR1F[d]	CAATGTGCCAAAGATGAGCG	4007–4026	47	1066	Pospieszny *et al.*, 2007
	ToT4b-R[c,e]	AGTTCCTTGATGAGCCCAATG	5053–5073	47		Budziszewska *et al.*, 2008
MP (RNA2)	ToTVB-R[d]	TTCCAGTAATGATCCAACCAAT	1055–1076	56	576	Van den Heuvel *et al.*, 2006
	ToTVB-D[e]	CCCATCATCACCCTCCTCTTCGTA	1608–1631	66		Van den Heuvel *et al.*, 2006
Vp35 (RNA2)	Vp35-D[d]	ACTGGAAAGGTACGGCATTGTGAA	1880–1903	51	694	This study
	ToTVCP-R[e]	GACCCGAAATTGCACCCATGCCGG	2551–2574	57		This study
Vp23 (RNA2)	TR2F[d]	GAAGGACGAAGAGCGACTG	3683–3701	48	573	Pospieszny *et al.*, 2007
	TR2R[e]	AAGGTAGGTATGCGTTTGC	4238–4256	44		Pospieszny *et al.*, 2007

[a]Nucleotide location refers to sequences of isolate Wal03 available in GenBank (Acc. Nos EU563947 and EU563947).
[b]In which primer was designed.
[c]Primer used in the opposite direction to which it was designed.
[d]Specific forward primers.
[e]Specific reverse primers.

www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html). Similarly, the amino acid substitution model that maximized the likelihood of the observation was inferred using the PROTTEST server (http://darwin.uvigo.es/software/prottest_server.html). Phylogenetic analyses based on the nucleotide and amino acid sequences were performed with MEGA (Molecular Evolutionary Genetics Analysis) version 3.1 (Kumar *et al.*, 2004) using a neighbour-joining algorithm. The statistical reliability of the constructed tree was assessed by the bootstrap method based on 10 000 pseudoreplicates. The comparative analyses were performed using sequences available in GenBank. The genetic distances between pairs of sequences and pairwise synonymous (d_S) and nonsynonymous (d_{NS}) substitutions were calculated according to the Kimura two-parameter method (Kimura, 1980) and the method of Pamilo, Bianchi and Li (Li, 1993; Pamilo & Bianchi, 1993), using MEGA. Confidence estimates for nonsynonymous (amino-acid-replacing) and synonymous (silent mutational) nucleotide substitutions were calculated by using the bootstrap method with 1000 replicates. To identify the specific amino acids under selective constraints, the differences between the nonsynonymous (d_{NS}) and synonymous (d_S) substitution rates were estimated for each position in the alignment using the fixed-effects likelihood (FEL) and internal fixed-effects likelihood (IFEL) method (Kosakovsky Pond & Frost, 2005; Kosakovsky Pond *et al.*, 2006), as implemented in HYPHY (http://www.datamonkey.org). A $d_{NS} - d_S$ value >0 is taken as evidence for positive selection, whereas a value <0 is a sign of negative selection.

Results

The genetic diversity of the Spanish isolates of ToTV was analysed by comparing the nucleotide and amino acid sequences of five selected genome fragments: Pro-Co and RdRp (RNA1) and MP, Vp23 and Vp35 (RNA2) as previously defined. In addition, three Hungarian isolates were also included in the comparative analyses. The genomic regions studied were selected by being representative of the two RNAs of the ToTV genome. Moreover, the fragments of the RdRp, Vp35 and Vp23, constituted regions of the ToTV genome with more amino acid substitutions (Budziszewska *et al.*, 2008).

The percentages of similarity/identity of the predicted amino acid sequences showed that the studied isolates were very similar to each other and also to the published ToTV type (PRI-ToTV0301) and Polish Wal03 isolates (Table 3). The identity/similarity among the amino acid predicted sequences ranged from 97·2 to 100%. The most similar genomic region studied was the RdRp followed by Pro-Co, and the highest differences were found in the RNA2 regions, mainly in Vp23. Isolate W4-07, collected in 2007 from weed species, presented the lowest identity/similarity in the MP compared with the published ToTV isolates, ToTV type and Wal03 (Table 3).

Phylogenetic analyses among the nucleotide and amino acid sequences of the 22 ToTV isolates studied were inferred for the five analysed genomic regions using the neighbour-joining method. It was clearly observed in the five regions studied that all the ToTV isolates clustered separately from ToMarV and ToANV, these being two isolates of a second species in the genus, *Torradovirus* as indicated Sanfaçon *et al.* (2009). The isolates studied were very similar and belonged to a single phylogenetic cluster. No significant grouping was observed according to the geographic origin or the collection year, except for the Gran Canaria isolates using the Vp35 (bootstrap 76%, Fig. 1a) and Vp23 (bootstrap 100%) regions. Interestingly, the grouping included the ToTV variant isolates from both tomato and weed species (*Solanum nigrum*, isolate W4-07). In addition, the Hungarian isolates did not group separately from the Spanish isolates in any of the studied regions. In general, there was no evidence of population subdivision, indicating a close genetic distance among the isolates. The phylogenetic analysis showed that the sequences of the different studied regions were much conserved, and that when the analysis was performed with the predicted amino acid sequences, all the ToTV isolates were grouped together and no differences were observed, even in the Vp23 and Vp35 sequences (Fig. 1b).

Genetic distances for each pair of isolates were estimated for each genomic region analysed by the two-parameter method of Kimura. The mean genetic distance ranged from 0·0044 to 0·117. Although all the studied regions of the genome for all the isolates presented similar genetic distances, the mean value for RdRp was half that of the other regions studied. The highest genetic distance was found for Vp23, although the differences among the RNA2 regions studied were minor (Table 4).

Pairwise genetic differences at the synonymous (d_S) and nonsynonymous (d_{NS}) nucleotide positions were estimated according to the method of Pamilo, Bianchi and Li (Table 4). The number of synonymous substitutions per synonymous site (d_S) was highest in Vp23 and lowest in RdRp. The nonsynonymous substitutions per nonsynonymous site (d_{NS}) were smaller than the d_S for all the studied regions. The ratio between nucleotide diversity values in nonsynonymous and synonymous positions (d_{NS}/d_S) provides an estimation of the degree and direction of the selective constraints operating in a coding region. Overall, the obtained values of the d_{NS}/d_S ratio for the five studied areas were markedly low (0·0000–0·2296, Table 4). The highest d_{NS}/d_S ratio was found in the MP region, indicating that this region was under looser functional constraints than the others studied. In the RdRp region, no nonsynonymous amino acid changes were observed; only some silent changes (synonymous) were present in the Spanish isolates (Table 4). In addition, nine negatively selected sites and just one site under positive selection were observed for all the proteins analysed ($P < 0·05$) using FEL analysis on the HYPHY server. In the case of RdRp, no selected site was observed; however, for the Pro-Co region, three negatively selected sites were

Table 3 Percentage of identity/similarity of amino acid sequences between the studied *Tomato torrado virus* (ToTV) isolates and the ToTV type isolate PRI-ToTV0301 and the Polish isolate Wal03

	RNA1								RNA2											
	Protease cofactor				RdRp				MP				Vp23				Vp35			
	PRI-ToTV0301[b]		Wal03[c]		PRI-ToTV0301[b]		Wal03[c]		PRI-ToTV0301[d]		Wal03[e]		PRI-ToTV0301[d]		Wal03[e]		PRI-ToTV0301[d]		Wal03[e]	
Isolate[a]	iden.[f]	simil.[f]	iden.	simil.	iden.	simil.	iden.	simil.	iden.	simil.	iden.	simil.	iden.	simil.	iden.	simil.	iden.	simil.	iden.	simil.
MUR-01	99·6	100·0	99·1	99·6	100·0	100·0	99·6	100·0	99·3	99·3	98·6	99·3	98·8	99·4	99·4	100·0	98·9	100·0	98·9	100·0
MUR-02	99·6	99·6	99·1	99·6	100·0	100·0	99·6	100·0	99·3	99·3	98·6	99·3	98·8	99·4	99·4	100·0	98·9	99·5	98·9	99·5
MUR-03	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
TEN-03	99·6	100·0	99·6	99·6	100·0	100·0	99·6	100·0	99·3	100·0	98·6	100·0	98·2	99·4	99·4	100·0	99·5	100·0	99·5	100·0
ALM-04	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	99·4	99·4	99·4	100·0	100·0	100·0	98·9	100·0
GNC-04	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	99·3	99·3	98·6	99·3	98·2	98·8	98·8	99·4	99·5	100·0	99·5	100·0
W3-04	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	98·4	99·5	98·4	99·5
GNC-05	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	97·6	98·2	98·2	98·8	98·9	100·0	98·9	100·0
MUR-05	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	98·8	99·4	100·0	98·9	99·5	98·9	99·5
GNC-06 (ToTV-CAN)	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·2	98·8	98·8	99·4	98·4	100·0	98·4	100·0
MUR-06	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
ALC-07	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
MUR-07	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
TEN-07	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	99·3	99·3	98·6	99·3	98·2	99·4	99·4	100·0	99·5	100·0	99·5	100·0
W4-07	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	97·9	98·6	97·2	98·6	98·2	98·8	98·8	99·4	99·5	100·0	99·5	100·0
BCN-08	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
GNC-08	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	99·3	99·3	98·6	99·3	98·2	98·8	98·8	99·4	99·5	100·0	99·5	100·0
MUR-08	99·6	99·6	99·1	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
MUR-09	99·6	99·6	99·1	99·1	100·0	100·0	99·6	100·0	99·3	100·0	98·6	100·0	98·2	98·8	98·8	99·4	99·5	100·0	99·5	100·0
ToTV-H1	99·6	100·0	99·1	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	99·5	100·0	99·5	100·0
ToTV-H2	99·6	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	98·8	99·4	99·4	100·0	98·9	99·5	98·9	99·5
ToTV-H3	100·0	100·0	99·6	99·6	100·0	100·0	99·6	100·0	100·0	100·0	99·3	100·0	97·6	99·4	99·4	100·0	99·5	100·0	99·5	100·0

[a]Spanish isolates collected from tomato were codified by three letters indicating geographic origin (ALC = Alicante, ALM = Almeria, BCN = Barcelona, GNC = Gran Canaria, MUR = Murcia, TEN = Tenerife), followed by the year of collection. Isolates from weeds have a 'W' prefix and Hungarian isolates have an 'H' suffix.

[b]ToTV-type isolate PRI-0301 GenBank Acc. No. DQ388879.

[c]Polish isolate Wal03 GenBank Acc. No. EU563948.

[d]ToTV-type isolate PRI-0301 GenBank Acc. No. DQ388880.

[e]Polish isolate Wal03 GenBank Acc. No. EU563947.

[f]Identity and similarity percentages of the deduced amino acid sequences calculated with MATGAT (Matrix Global Alignment Tool) version 2.01 (Campanella *et al.*, 2003).

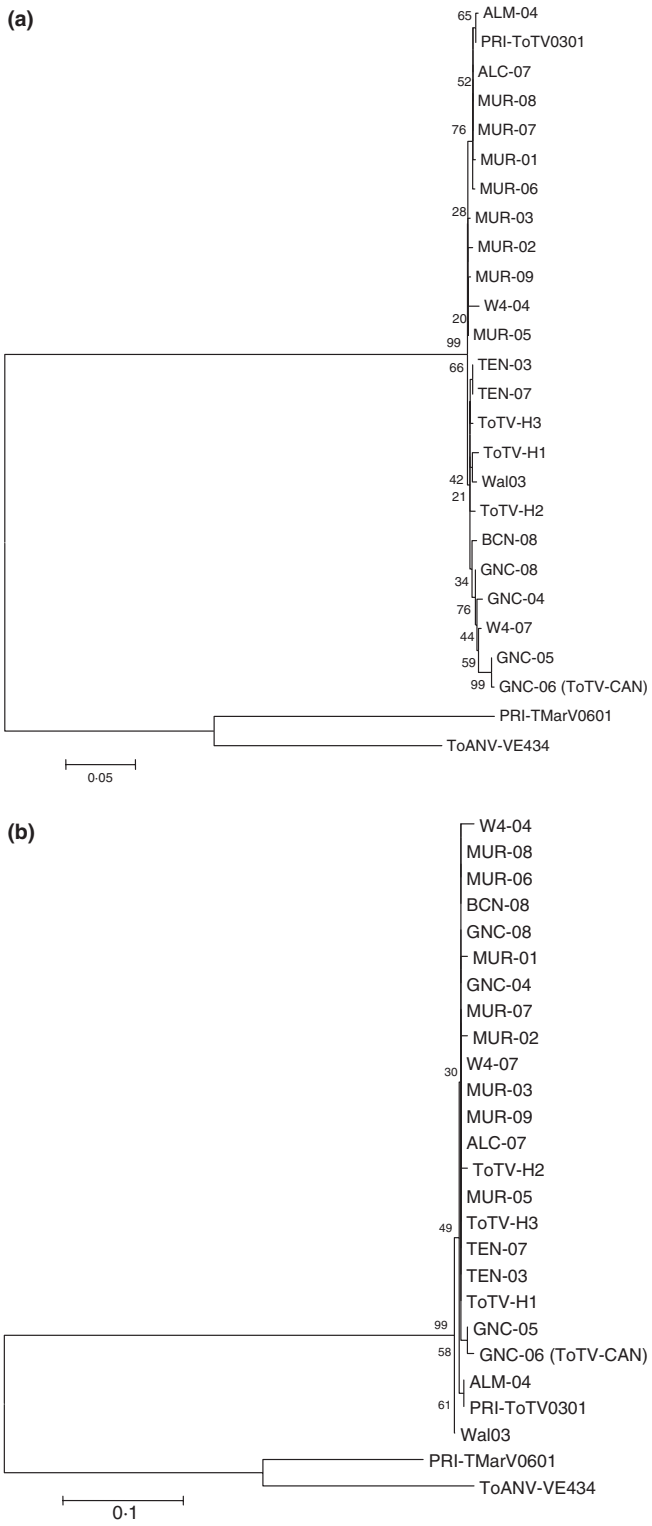

Figure 1 Representative phylogenetic analysis showing the relationship of the obtained nucleotide (a) and amino acid (b) sequences of CP subunit Vp35 of 19 Spanish and three Hungarian *Tomato torrado virus* (ToTV) isolates, with other isolates and related viruses. Neighbour-joining phylogenetic trees were obtained using MEGA version 3.1. (Kumar *et al.*, 2004). The statistical reliability of the constructed trees was assessed by the bootstrap method based on 10 000 pseudoreplicates. Numbers above nodes indicate percentages of bootstrap replicates which supported branching. Scale bars represent genetic distance of 0·05 (a) and 0·1 (b). PRI-ToMarV0601 GenBank Acc. Nos NC010987 and NC010988; ToANV-VE434 GenBank Acc. Nos EF063641 and EF063642. All other Accession Nos can be found in text.

Table 4 Average number of nucleotide substitutions among 19 Spanish and three Hungarian isolates of *Tomato torrado virus* (ToTV), the ToTV type isolate PRI-ToTV0301 and the Polish isolate Wal03 for five genomic regions

Genomic region	d^a	$d_{NS}{}^a$	$d_S{}^a$	$d_{NS}/d_S{}^a$
Pro-Co (RNA1)	0·0080 ± 0·0013	0·0017 ± 0·0008	0·0219 ± 0·0040	0·0776
RdRp (RNA1)	0·0044 ± 0·0011	0·000 ± 0·0000	0·0115 ± 0·0028	0·0000
MP (RNA2)	0·0102 ± 0·0017	0·0048 ± 0·0016	0·0209 ± 0·0043	0·2296
Vp23 (RNA 2)	0·0117 ± 0·0022	0·0033 ± 0·0014	0·0314 ± 0·0075	0·1051
Vp35 (RNA2)	0·0105 ± 0·0021	0·0024 ± 0·0008	0·0281 ± 0·0060	0·0854

[a]d = nucleotide diversity according to Kimuras's two-parameter method; d_{NS} = average number of nonsynonymous substitutions per nonsynonymous site; d_S = average number of synonymous substitutions per synonymous site (Pamilo & Bianchi, 1993); d_{NS}/d_S = average ratio between nonsynonymous and synonymous substitutions for each pair of comparisons.

obtained (S74, S212 and T301). For the MP, three sites were also detected under negative selection (I148, G156 and G244). Three negatively selected sites were observed for the CP, one for the Vp23 (E1151) and two for the Vp35 (R469 and A581). Using IFEL analysis to characterize the selected sites in the internal branches of the tree, the only case of positive selection was detected in the Pro-Co (V-I 291). The position of the negative and positive selection sites refers to the amino acid complete sequence of the polyprotein in each RNA.

Discussion

Plant RNA viruses are believed to have a great potential for genetic variation because their replication process is error-prone, since no proofreading correction mechanism has been associated with their RdRp (Holland *et al.*, 1982). Nonetheless, the large mutability need not result in high genetic variability, and other factors such as natural selection or bottlenecks can reduce the genetic variation of a virus (Roosnick, 1997). Studies of the variability and changes in the genetic structure of plant virus populations are an important aspect of plant pathology, and may even be highly relevant for the development of strategies to control virus-induced diseases (García-Arenal *et al.*, 2001).

This work analysed the genetic diversity of five different regions of the genome with different functions, from 19 ToTV Spanish isolates collected in the main tomato-growing areas over a 9-year period. Those isolates were also compared with three Hungarian isolates collected over a 2-year period and the available sequences of ToTV isolates published in GenBank. A phylogenetic analysis of the five genome areas studied revealed that all the ToTV isolates studied clustered together and separately from ToMarV and ToANV, two isolates of a second member of the new genus *Torradovirus*, as observed in previous studies (Verbeek *et al.*, 2007b; Budziszewska *et al.*, 2008; Pospieszny *et al.*, 2010). In this sense, the identity/similarity percentages among all the analysed ToTV isolates (including the ToTV variants from Spain, Poland and Hungary) ranged between 97 and 100% for all the analysed proteins, but were lower than 85% when compared with ToMarV and ToANV (Budziszewska *et al.*, 2008). ToTV and ToMarV (ToANV is proposed as a different isolate of ToMarV) have been proposed to constitute a new genus, *Torradovirus*. This new genus has also been proposed to belong to the new family, Secoviridae, which will also include other genera such as *Sequivirus*, *Waikavirus*, *Comovirus*, *Fabavirus*, *Nepovirus*, *Cheravirus*, *Sadwavirus* and other tentative viruses of this last genus. All these genera present common properties, such as particle structure (icosahedral with pseudo $T = 3$ symmetry), a positive-strand RNA genome that has a polyprotein expression strategy, a common replication block including type-III helicase, 3C-like cysteine proteinase, type I polymerase, and they clustered together as a single branch in the Pro-Pol dendrogram when compared to other picornaviruses (Sanfaçon *et al.*, 2009).

The genetic diversity within the ToTV isolates was very low (0·0044–0·0117, Table 4), but similar to reported values for other viruses (Rubio *et al.*, 2001; Janssen *et al.*, 2007). There was no evidence of any temporal differentiation on a local scale of the Spanish population as occurs in *Tomato yellow leaf curl Sardinia virus* (TYLCSV, Sánchez-Campos *et al.*, 2002), or *Cucurbit yellow stunting disorder virus* (CYSDV, Marco & Aranda, 2005), or of any geographical differentiation in the studied isolates, in general, as in the Spanish isolates of CYSDV (Rubio *et al.*, 2001), *Tobacco mild green mosaic virus* (TMGMV, Fraile *et al.*, 1996) or *Watermelon mosaic virus* (WMV, Moreno *et al.*, 2004). Unlike the rest of the amino acid sequences analysed, the two CP subunits grouped all the Gran Canaria isolates in the same cluster, including one isolate characterized from the weed species *S. nigrum* from Tenerife, an island close to Gran Canaria. The observation that the Gran Canaria variants clustered together with the weed variant from Tenerife (W4-07) could suggest a common origin for all these isolates (GNC-04, GNC-05, GNC-06, GNC-08 and W4-07), different from that of the rest of the ToTV isolates, including the other Tenerife isolates. Geographic grouping correlated to variation in the CP gene has also been reported for other viruses, such as Iranian WMV isolates (Sharifi *et al.*, 2008) and Tunisian *Zucchini yellow mosaic virus* (ZYMV) isolates (Yakoubi *et al.*, 2008).

All the ToTV genome sequences analysed were characterized by high genetic stability that could be attributed to the selection to maintain the functional integrity of the viral genome (García-Arenal *et al.*, 2001). The degree of

negative selection was estimated by the ratio between the nucleotide diversities in the nonsynonymous and synonymous positions (d_{NS}/d_S), which indicates the amount of variation in the nucleic acid that results in the variation of the encoded protein. For the ToTV population studied herein, this ratio <1 (i.e. the nonsynonymous lower than the synonymous substitution rate) for all the coding regions, thus suggesting that all proteins are under high negative selective constraints. Interestingly, the MP genomic region presented a slightly higher value than the two CP regions. However, this variation was not reported to correlate with the function of the encoded protein, and varied largely according to the gene and the virus (García-Arenal *et al.*, 2001).

In addition, ecological conditions affecting the virus, vector, host plant and their interactions, such as dominant selection pressure, may contribute to generic stability (García-Arenal *et al.*, 2001), as observed in other whitefly-transmitted viruses such as *Cucumber vein yellowing virus* (CVYV) (Janssen *et al.*, 2007) or CYSDV (Rubio *et al.*, 2001). The overall variability observed suggests a genetic stability that could be explained, at least in part, by the ecological bottleneck in which host plants and vectors are continuously maintained (Fraile *et al.*, 1996; Sánchez-Campos *et al.*, 2002). On the other hand, a viral population which shows higher genetic variability is normally considered to be older, although this may or may not always be consistent (Wei *et al.*, 2009). In this sense, the adaptation of a plant virus to a new resistant host or species implies specific molecular changes that confer some advantages to the new host (Jenner *et al.*, 2002; Wallis *et al.*, 2007). A recent example of fast host-specific adaptation is the convergent evolution of *Pelargonium flower break virus* (PFBV) populations adapted to *Chenopodium quinoa* (Rico *et al.*, 2006). In this example, PFBV incorporated five specific noncontiguous amino acids in the CP when it was maintained in the *C. quinoa* host. ToTV was recently detected in tomato crops and some weed species collected in tomato-growing greenhouses (Alfaro-Fernández *et al.*, 2008) and has not spread widely in many countries. Remarkably, this virus seems to be adapted mainly to Solanaceae species (Pospieszny *et al.*, 2010), although it naturally infects some weed species of different families (Alfaro-Fernández *et al.*, 2008). The high sequence stability observed in the coding regions could be interpreted as ToTV having already adapted to the tomato host. However, a significant grouping of the Gran Canaria isolates together with one from a solanaceous weed was observed, in which some residues were specific for all the group components, such as A1051 in Vp23, or others like G302 (W4-07 and GCN-08) and Y303 (W4-07, GCN-08, GCN-04) in the MP or the positively selected site V-I 291 in the Pro-Co that contains a specific I (isoleucine) amino acid for the W4-07 and GCN-08 isolates. This observation raises two possibilities: (i) that a putative ToTV from a weed host could evolve in the new tomato host, or vice versa, and (ii) if this were the case, then we can speculate a putative origin of the recent new ToTV in tomato crops. This possibility is currently being investigated by analysing the stability of the ToTV progeny in the different weed and/or tomato hosts concerned. In addition, it would be very interesting to study newly emerging populations of this virus which could become a problem for tomato production worldwide.

Acknowledgements

This work was supported by grants AGL2005-06682-C03-01 from the Spanish Ministry of Education and Science (MEC). We thank the Instituto Agroforestal Mediterráneo (UPV, Valencia) for fellowship support to AAF. We also thank Dr M. Juárez (Universidad Miguel Hernández, Orihuela) and Dr A. Lacasa (IMIDA, Murcia) for their advice and assistance in the field surveys. We wish to also thank Bese Gabor (Csongrád Megyei MgSzH) and A. Forray (Floraton Kft.) for providing the Hungarian isolates of ToTV included in this assay.

References

Alfaro-Fernández A, Córdoba-Selles MC, Cebrián MC *et al.*, 2007. First report of Tomato torrado virus in tomato in the Canary Islands, Spain. *Plant Disease* **91**, 1060.

Alfaro-Fernández A, Córdoba-Sellés MC, Cebrián MC *et al.*, 2008. First report of Tomato torrado virus on weed hosts in Spain. *Plant Disease* **92**, 831.

Alfaro-Fernández A, Bese G, Córdoba-Sellés MC *et al.*, 2009. First report of Tomato torrado virus (ToTV) infecting tomato in Hungary. *Plant Disease* **93**, 554.

Alfaro-Fernández A, Medina V, Córdoba-Sellés MC *et al.*, 2010. Ultrastructural aspects of tomato leaves infected by Tomato torrado virus (ToTV) and co-infected by other viruses. *Plant Pathology* **59**, 231–9.

Amari K, González-Ibeas D, Gómez P *et al.*, 2008. Tomato torrado virus is transmitted by *Bemisia tabaci* and infects pepper and eggplant in addition to tomato. *Plant Disease* **92**, 1139.

Budziszewska M, Obrepalska-Steplowska A, Wieczorek P, Pospieszny H, 2008. The nucleotide sequence of a Polish isolate of Tomato torrado virus. *Virus Genes* **37**, 400–6.

Campanella JJ, Bitincka L, Smalley J, 2003. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* **4**, 29–33.

EPPO, 2009. *EPPO Alert List*. Paris, France: European and Mediterranean Plant Protection Organization.

Fraile A, Malpica JM, Aranda MA, Rodriguez-Cerezo E, García-Arenal F, 1996. Genetic diversity in *Tobacco mild green mosaic tobamovirus* infecting wild plant *Nicotiana glauca*. *Virology* **223**, 148–55.

García-Arenal F, Fraile A, Malpica JM, 2001. Variability and genetic structure of plant virus populations. *Annual Review of Phytopathology* **39**, 157–86.

Herrera-Vásquez JA, Alfaro-Fernández A, Córdoba-Sellés MC, Cebrián MC, Font MI, Jordá C, 2009. First report of Tomato torrado virus infecting tomato in single and mixed infections with *Cucumber mosaic virus* in Panama. *Plant Disease* **93**, 198.

Holland JJ, Spindler K, Horodyski F, Grabau E, Nochol S, VandePol S, 1982. Rapid evolution of RNA genomes. *Science* **215**, 1577–85.

IPPC, 2008. *Detection of Tomato torrado Virus in South Australia*. Canberra, Australia: International Pest Protection Convention: Official Pest Report No. AU-14/1.

Janssen D, Velasco L, Martín G, Segundo E, Cuadrado IM, 2007. Low diversity among *Cucumber vein yellowing virus* isolates from Spain. *Virus Genes* **34**, 367–71.

Jenner CE, Wang X, Ponz F, Walsh JA, 2002. A fitness cost for *Turnip mosaic virus* to overcome host resistance. *Virus Research* **86**, 1–6.

Kimura M, 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–20.

Kosakovsky Pond SL, Frost SDW, 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution* **22**, 1208–22.

Kosakovsky Pond SL, Frost SDW, Grossman Z, Gravenor MB, Richman DD, Brown AJL, 2006. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. *PLoS Computational Biology* **2**, 531–8.

Kumar S, Tamura K, Nei M, 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5**, 150–63.

Li W-H, 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *Journal of Molecular Evolution* **36**, 96–9.

MacKenzie DJ, Mclean MA, Mukerji S, Green M, 1997. Improved RNA extraction from woody plants for the detection of viral pathogens by reverse transcription-polymerase chain reaction. *Plant Disease* **81**, 222–6.

Marco CF, Aranda MA, 2005. Genetic diversity of a natural population of *Cucurbit yellow stunting disorder virus*. *Journal of General Virology* **86**, 815–22.

Moreno IM, Malpica JM, Díaz-Pendón JA, Moriones E, Fraile A, García-Arenal F, 2004. Variability and genetic structure of the population of *Watermelon mosaic virus* infecting melon in Spain. *Virology* **318**, 451–60.

Pamilo P, Bianchi NO, 1993. Evolution of the Zfx and Zfy genes: rates and independence between the genes. *Molecular Biology and Evolution* **10**, 271–81.

Pospieszny H, Borodynko N, Obrepalska-Steplowska A, Hasiów B, 2007. First report of Tomato torrado virus in Poland. *Plant Disease* **91**, 1364.

Pospieszny H, Budziszewska M, Hasiów-Jaroszewska B, Obrepalska-Steplowska A, Borodynko N, 2010. Biological and molecular characterization of Polish isolates of Tomato torrado virus. *Journal of Phytopathology* **158**, 56–62.

Rico P, Ivars P, Elena SF, Hernández C, 2006. Insights into the selective pressures restricting *Pelargonium flower break virus* genome variability: evidence for host adaptation. *Journal of Virology* **80**, 8124–32.

Roosnick MJ, 1997. Mechanisms of plant virus evolution. *Annual Review of Phytopathology* **35**, 479–89.

Rubio L, Abou-Jawdah Y, Lin H-X, Falk BW, 2001. Geographically distant isolates of the crinivirus *Cucurbit yellow stunting disorder virus* show very low genetic diversity in the coat protein gene. *Journal of General Virology* **82**, 929–33.

Sánchez-Campos S, Díaz JA, Monci F *et al.*, 2002. High genetic stability of the begomovirus *Tomato yellow leaf curl virus* in southern Spain over an 8-year period. *Phytopathology* **92**, 842–9.

Sánchez-Navarro JA, Aparicio F, Rowhani A, Pallás V, 1998. Comparative analysis of ELISA, non-radioactive molecular hybridization and PCR for the detection of *Prunus necrotic ringspot virus* in herbaceous and *Prunus* hosts. *Plant Pathology* **47**, 780–6.

Sanfaçon H, Wellink J, Le Gall O, Karasev A, Van der Vlugt R, Wetzel T, 2009. Secoviridae: a proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned *Cheravirus* and *Sadwavirus*, and the proposed genus Torradovirus. *Archives of Virology* **154**, 899–907.

Sharifi M, Massumi H, Heydarnejad J, Pour AH, Shaabanian M, Rahimian H, 2008. Analysis of the biological and molecular variability of *Watermelon mosaic virus* isolates from Iran. *Virus Genes* **37**, 304–13.

Turina M, Ricker MD, Lenzi R, Masenga V, Ciuffo M, 2007. A severe disease of tomato in the Culiacan area (Sinaloa, Mexico) is caused by a new picorna-like viral species. *Plant Disease* **91**, 932–41.

Van den Heuvel JF, Maris PC, Verbeek M, Dullemans AM, Van der Vlugt RA, 2006. *Plant Virus Designated Tomato torrado Virus*. Geneva, Switzerland: World Intellectual Property Organization: Patent no. WO2006085749.

Verbeek M, Dullemans AM, Van den Heuvel JFJ, Maris PC, Van der Vlugt RAA, 2007a. Identification and characterization of Tomato torrado virus, a new plant picorna-like virus from tomato. *Archives of Virology* **152**, 881–990.

Verbeek M, Dullemans AM, Van den Heuvel JFJ, Maris PC, Van der Vlugt RAA, 2007b. Tomato marchitez virus, a new plant picorna-like virus from tomato related to tomado torrado virus. *Archives of Virology* **153**, 127–34.

Verdin E, Gognslons P, Wipf-Scheibel C *et al.*, 2009. First report of Tomato torrado virus in tomato crops in France. *Plant Disease* **93**, 1352.

Wallis CM, Stone AL, Sherman DJ, Damsteegt VD, Gildow FE, Schneider WL, 2007. Adaptation of *Plum pox virus* to a herbaceous host (*Pisum sativum*) following serial passages. *Journal of General Virology* **88**, 2839–45.

Wei T-Y, Yang J-G, Liao F-L *et al.*, 2009. Genetic diversiy and population structure of *Rice stripe virus* in China. *Journal of General Virology* **90**, 1025–34.

Yakoubi S, Desbiez C, Fakhfakh H *et al.*, 2008. Molecular, biological and serological variability of *Zucchini yellow mosaic virus* in Tunisia. *Plant Pathology* **57**, 1146–54.