A novel distinct genetic variant of tomato torrado virus with substantially shorter RNA1-specific 3'untranslated region (3'UTR)

Marta Budziszewska
Institute of Plant Protection National Research Institute: Instytut Ochrony Roslin
Przemysław Wieczorek (p.wieczorek@iorpib.poznan.pl)
Institute of Plant Protection National Research Institute: Instytut Ochrony Roslin
https://orcid.org/0000-0002-3306-9763
Aleksandra Obrępalska-Stęplowska
Institute of Plant Protection National Research Institute: Instytut Ochrony Roslin

Research Article

Keywords: tomato torrado virus, genetic variability, 3'untranslated regions, tandem repeats

Posted Date: May 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-502135/v1

Version of Record: A version of this preprint was published at Plants on November 13th, 2021. See the published version at https://doi.org/10.3390/plants10112454.

Abstract

Tomato torrado virus (ToTV) induces severe systemic necrosis in *Solanum lycopersicum*. This work aimed at describing the genetic variability of necrosis-inducing ToTV-Wal'17 collected in the 2017 year, derived from the ToTV-Wal'03 after long-term passages in plant. Sequence analyses of the ToTV-Wal`17 indicated twenty-eight single nucleotide substitutions in coding sequence of both RNAs, twelve of which resulted in amino acid changes in viral polyproteins. Moreover the sequencing data revealed that the 3'UTR of ToTV-Wal'17 RNA1 was 394 nts shorter in comparison to Wal'03.

The performed sequence analyses pointed that 3'UTR of RNA1of ToTV-Wal'17 is the most divergent across all previously described European isolates.

Introduction

Tomato torrado virus (ToTV) is a type member of the *Torradovirus* genus in the family *Secoviridae* [1, 2]. In its natural environment, ToTV infects *Solanum lycopersicum* causing `torrado` disease, manifested by burn-like severe necrosis of stems, leaves, and fruits [3–5]. In the field or greenhouses, the virus is transmitted by whiteflies [3, 6]. Tomato torrado virus was identified in Spain [7], Hungary [8], Poland [6], Canary Islands [9], France [10], Panama [11], Italy [12], Australia [13], Colombia [14], and Morocco [15]. Because of its expansion and severity, in the years 2009–2013, the ToTV was included in EPPO Alert List (https://gd.eppo.int/taxon/TOTV00/documents). Recently, the ToTV was found infecting tomato crops in new locations: South Africa [16] and Serbia [17], indicating that the virus should still be considered as serious threat to tomato. Under experimental conditions, ToTV can be mechanically inoculated to, among others, *S. lycopersicum* or *N. benthamiana* [4, 18], a well-known model for plant-pathogen interactions [19].

The ToTV has a bipartite genome with RNA1 (7802–7814 nucleotides, excluding polyA tail) and RNA2 (5390 nucleotides, excluding polyA tail). Coding sequences in RNA1 and RNA2 are flanked with untranslated regions (UTRs). The 5'UTRs of RNA1 and RNA2 of ToTV are composed of 107 nts and 181 nts, respectively, and remain genetically conserved between described ToTV isolates. Compared to the known RNA plant viruses, ToTV, like other torradoviruses, has very long 3'UTRs in both genomic strands [3, 4]. The 3'UTR in RNA1 (1230 nts) has a variable region (VR) of 241 nts followed by 989 nts of the conserved region (CR); the 3'UTR in ToTV RNA2 is shorter (1106 nts) and begins with 103 nts followed by CR sequence having 98% similarity with the corresponding region from RNA1. Noteworthy, within the CR in both RNA1 and RNA2, two direct repeats were identified: D1a/D1b and D2a/D2b arranged in the following order: D1b-D2a-D1a-D2b [20]. Notably, described full-length sequences of the Polish ToTV Kra and Ros isolates revealed remarkable length heterogeneity in VR that arose from nucleotide deletions in this region [20, 21].

RNA viruses are known to have a high mutation rate because virus RNA polymerase lacks proofreading activity. Thus, during each replication cycle, a pool of mutated RNA species is generated, and in particular

environmental conditions, only the optimal RNA master sequence will preferentially accumulate. This is one of the well-known mechanisms of the emergence of virus genetic variability [22].

In this work, we assessed the genetic variability of new ToTV isolate, the ToTV-Wal'17 originated from the previously described ToTV-Wal'03, causing severe necrosis on tomato and leaf malformations on *N. benthamiana*. The ToTV-Wal'03 was mechanically passaged spontaneously over 13 years and isolated from *S. lycopersicum* in 2017 for comparative genetic analyses. Here we showed that apart from the several new point mutations identified in the ToTV-Wal'17 genome, the ToTV-Wal'17 has significantly shorter RNA1 3'UTR and, as such, remained still highly infectious to tomato. Compared to previously described ToTV isolates, the mentioned RNA1 3'UTR of ToTV-Wal'17 was described to be shorter of 394 nts in the CR region. This, in turn, changed significantly the conserved stretch of the D1b-D2a-D1a-D2b sequences. The performed analyses pointed that in the context of 3'UTR in RNA1, the ToTV-Wal'17 is the most diverse across all previously described European isolates of ToTV.

Materials And Methods

The ToTV-Wal`03 [4] was used as a virus source. The virus was maintained and propagated on *N. benthamiana* or alternately in *S. lycopersicum* plants, periodically collected, stored at -20□C, and used as the source of virus inocula for thirteen years [5]. In 2017 from the infected plants (*S. lycopersicum*) virus RNA was isolated and named hereafter ToTV-Wal'17.

The Total RNA was extracted from diseased tomato plants using TriReagent [19] and cDNA was synthesised in the presence of the asTo2C_pJL_RV primer [18] and SuperScript IV Reverse Transcriptase (Thermo Scientific) following the manufacturer's instructions. The cDNA, covering the full-length RNA1 and RNA2, was amplified utilizing polymerase chain reaction with CloneAmp HiFi PCR Premix (Takara) and primers asTo1A_pJL_FW/asTo2C_pJL_RV (RNA1) and asTo2A_pJL_FW/asTo2C_pJL_RV (RNA2) [18]. The cDNA copies were fused with pJL89 plasmid (NEBuilder HiFi DNA Assembly Master Mix, NEB) and cloned in *E. coli* Stellar competent cells (Takara). Then, the recombined plasmids were Sanger sequenced to generate overlapping contigs. The retrieved contig sequences were assembled using BioEdit software [23], and the obtained consensus sequences of RNA1 and RNA2 were analysed.

The sequences of Wal'17 isolate were aligned with genomic sequences of Wal`03 (EU563948, EU563947) using Basic Local Alignment Search Tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Additionally, the multiple nucleotides and amino acid sequence analyses of all known ToTVs isolates were conducted using MEGAX software [24], followed by the phylogenetic studies [13], a maximum-likelihood algorithm, and 1000 bootstrap value.

Results And Discussion

The obtained complete sequences of RNA1 and RNA2 of ToTV-Wal'17 have been deposited in the NCBI database (MW729382 and MW729383). The lengths of RNA1 and RNA2 of the ToTV-Wal'17 were

calculated and showed to have 7419 and 5390 nts (excluding polyA tail), respectively, indicating that RNA1 of Wal'17 was 394 nts shorter than the corresponding RNA strand from ToTV-Wal'03. The performed comparative sequence analyses of RNA1 of ToTV-Wal'17 showed 99% (in a range of 1-6826 nts), and 77% (6899–7419 nts) sequence identity within the corresponding RNA of ToTV-Wal'03 (EU563948), respectively. On the other hand, the entire RNA2 had 99% sequence identity with the corresponding RNA of Wal'03 (EU563947). All indicated nucleotide substitutions in both RNAs of Wal'17 were listed in Table 1.

Table 1 The nucleotide (nt) and amino acid (aa) substitutions in particular functional fragments of RNA1 and RNA2 of Wal'17 (MW729382, MW729383) in comparison to original Wal`03 (EU563948, EU563947) sequences.

Prot - protease; Hel - helicase, RdRp – RNA dependent polymerase, UTR- untranslated region, ORF- open reading frame, 3A- protein involved in virus movement, 11K- domain

RNA strand	Silent mutations		Missense mutations	
	Nucleotide change	Functional domain	Nucleotide change (amino acid change)	Functional domain
RNA1	830G > A	Prot	346T > A (V80D)	11K
	1574C > T	Hel		
	1964C > T	Hel/Prot	1905G > A (D600N); 2058C > T (Y651H); 2242A > G (N712S); 2504G > C (K799N)	Hel
	2216G > A;			
	2612T > C			
	3500A > G;	Prot/RdRP	3436T > C (V1110A)	Prot/RdRP
	3734C > T;			
	4124T > C	RdRP		
	4256T > C			
	4496C > T			
	5603G > A	RdRP/3`UTR		
	5810C > T		4762G > A (R1552K)	RdRP
	6101C > T		5843G > C (Q1912H)	RdRP /3`UTR
	6473C > T			
RNA2	632G > A	ORF1	1863A > C (N388H); 1865T > C (N388H); 1866C > A (H389N); 2001C > T (L434F)	3A
	2573T > C	Vp35		

Analysis of 5'untranslated regions (5'UTRs)

Analysis of ToTV-Wal'17 RNA1 and RNA2 5'UTRs showed that the regions had the same length and nucleotide sequences as the 5'UTRs from the ToTV-Wal'03 [4].

Analysis of coding sequences (CDS)

The coding sequences of ToTV-Wal'17 had the same length as in ToTV-Wal'03 and comprised 6476 and 4116 nts in RNA1 and RNA2, respectively. In RNA1, twenty-two nucleotide substitutions were identified, of which eight of them changed the amino acid. Concerning RNA2, six nucleotide substitutions were indicated in CDS, and four of them resulted in amino acids substitution. All the described mutations in CDS were listed in Table 1.

Analysis of 3'untranslated regions (3'UTRs)

The RNA1 and RNA2 3'UTRs of ToTV-Wal'17 had 836 and 1092 nts, respectively. In general, in all ToTV isolates sequenced so far, the 3'UTR in RNA1 is longer than in RNA2. Surprisingly, contrary to ToTV-Wal'03, the 3'UTR in RNA1 of ToTV-Wal'17 was 256 nts shorter than that in RNA2. Moreover, by comparing sequence length, it was shown that 3'UTR in RNA1 of ToTV-Wal'17 was 394 nts shorter (32%) than the corresponding sequence from ToTV-Wal'03. On the other hand, the RNA2 3'UTR of ToTV-Wal'17 was the same in length as the analogous sequence from ToTV-Wal'03.

Previously it was shown that 3'UTR in ToTV RNAs were predicted to form structured regions [20]. Additionally, RNA1 and RNA2 3'UTRs of European isolates of ToTV are generally long (over 1000 nucleotides), and they are organised according to the following scheme: stop codon, followed by variable region (D-VR, 241 and 103 nts in RNA1 and RNA2, respectively), then conserved region (D-CR, 989 nts in RNA1 and RNA2) with terminal polyA tail. Moreover, in the CR in the 3'UTRs direct repeats were identified and arranged in the following repetitive motifs: D1b – D2a – D1a – D2b [20].

Concerning the ToTV-Wal'17, sequence comparison of its RNA1 3'UTR indicated a short variable region D-VR was preserved in the viral RNA. However, the D-CR region in RNA1 shared only 77% identity with the corresponding D-CR region from ToTV-Wal'03. Due to the high sequence variability, the repeated sequenced D1b – D2a – D1a – D2b were not identified. Notably, within the D-CR of the 3'UTR of RNA1 in ToTV-Wal'17, three other repetitive sequences of 70–71 nts in length (designed as CR_{RNA1}-R1 (position 6807–6877), CR_{RNA1}-R2 (6899–6969), CR_{RNA1}-R3 (7171–7240)) were identified (Fig. 1). The three mentioned repetitive sequences share 90–96% identity with the D2b region ToTV-Wal'03 [21] [20].

The RNA2 3'UTR of ToTV-Wal'17 shares 99% identity with the corresponding region of ToTV-Wal'03 and in the CR region, the repetitive motifs D1b – D2a – D1a – D2b were still present.

Additionally, performed phylogenetic studies based on the 3'UTR of RNA1 and RNA2 confirmed the results from comparative sequences studies mentioned above. In the dendrogram generated based on

3'UTR RNA1, we showed that the ToTV-Wal'17 isolate clustered with the recently reported South African isolate, and they both are phylogenetically distant from the other Polish and Spanish isolates of ToTV. The 3'UTR of RNA2 sequence analysis showed that the ToTV-Wal'17 clustered together with the ToTV-Wal'03 and had a short evolutionary distance with other Polish and Spanish isolates (Fig. 2).

Here we characterised new mutations (silent and missense) accumulated in ToTV-Wal'17 genome after mechanical passages from host-to-host over time. We observed that the performed mechanical transmission of the virus did not affect its pathogenicity: neither attenuated nor boosted virus severity on *S. lycopersicum* or *N. benthamiana*.

The high mutation rate of plant RNA viruses is directed by virus RNA polymerase without proofreading activity [22]. However, environmental pressure favours such viral RNA molecules (genomes) that can efficiently replicate, infect and spread in the host [22]. Therefore, it is not surprising that a few new mutations changing amino acid context in virus polyproteins were described in the coding region of the ToTV-Wal'17 genome after thirteen years of mechanical passages through the same hosts (tomato and *N. benthamiana*). However, what might be surprising is a new genetic variant of virus RNA1 with substantially shortened 3'UTR. So significantly shortened sequence and its high variability have not been described so far for other European isolates of ToTV. The comparative sequence studies of 3'UTRs of RNA1 of ToTV isolates deposited in GenBank between 2007–2014 showed similar lengths; only in Polish isolates Kra and Ros ToTV the additional deletion variants were observed.

On the other hand, the new ToTV isolate reported in 2020 by Moodley in comparison to the present Wal'17 ToTV indicates a further reduction in this region of the viral genome. However, in the Wal'17 isolate, the deletions occurred in the D-CR region, which is almost identical in both RNA strands, suggesting that this region might be, to some extent, redundant. Thus, the question arises, what was the reason of significant RNA1 3'UTR shortening in these two geographically distant isolates, including the fact that Wal'17 evolved from a much longer prototype variant of Wal'03.

Our previous studies described the genetic variability of 3'UTR in RNA1 of ToTV-Kra and ToTV-Ros, associated with diverse deletions concerning only the D-VR region [20, 21]. However, in the mentioned ToTV isolates, the D-CR regions remained relatively unchanged, with preserved repetitive motifs possibly involved in virus adaptation to hosts. It was discussed then that this sequence conservation in D-CR of RNA1 and RNA2 3'UTR fixated the presence of predicted secondary (and possibly tertiary) intramolecular interactions [20] that were anticipated to be created within the RNAs. Budziszewska *et al.* discussed the phenomenon of genetic heterogeneity in RNA1 of Polish ToTV isolates, suggesting that the variable in lengths variants of 3'UTRs may have a potential role in the virus replication process and/or transmission by insect vector [20]. The sequencing data and greenhouse observation obtained in a present study suggest that the regulatory elements (involved in virus replication) might be located outwardly in the D-CR region. The mechanism leading to the shortening of RNA1 3'UTR remains unknown. It might have been spontaneous or could have been forced by virus adaptability to plant host due to long-term passaging. Further experiments explaining the effect of identified variability in RNA1 and the function of

characteristic direct repeats in ToTV 3'UTRs RNAs are necessary to be performed to understand the ToTV evolution concerning plant-virus-vector interaction.

Declarations

Founding: This work was financially supported by grants from Polish National Science Centre to MB (2016/21/D/NZ9/02468) and PW (2016/21/D/NZ9/02478).

Conflicts of interest: The authors declare no conflict of interest.

References

1. Sanfaçon H, Wellink J, Le Gall O, et al (2009) Secoviridae: a proposed family of plant viruses within the order Picornavirales that combines the families Sequiviridae and Comoviridae, the unassigned genera Cheravirus and Sadwavirus, and the proposed genus Torradovirus. Arch Virol 154:899–907
2. Verbeek M, Dullemans AM, Maris P, et al (2010) The genus Torradovirus, a new plant virus genus. In: International Advances in Plant Virology, Joint meeting of Association of Applied Biologists & Dutch Circle of Plant Virologists
3. Verbeek M, Dullemans AM, den Heuvel J, et al (2007) Identification and characterisation of tomato torrado virus, a new plant picorna-like virus from tomato. Arch Virol 152:881–890
4. Budziszewska M, Obrepalska-Steplowska A, Wieczorek P, Pospieszny H (2008) The nucleotide sequence of a Polish isolate of Tomato torrado virus. Virus Genes 37:400–406
5. Pospieszny H, Budziszewska M, Hasiów-Jaroszewska B, et al (2010) Biological and molecular characterization of Polish isolates of Tomato torrado virus. J Phytopathol 158:56–62
6. Pospieszny H, Borodynko N, Obrepalska-Steplowska A, Hasiow B (2007) The first report of Tomato torrado virus in Poland. Plant Dis 91:1364
7. Alfaro-Fernández A, Córdoba-Sellés C, Cebrián MC, et al (2008) First report of Tomato torrado virus on weed hosts in Spain. Plant Dis 92:831
8. Alfaro-Fernández A, Bese G, Córdoba-Sellés C, et al (2009) First report of tomato torrado virus infecting tomato in Hungary. Plant Dis 93:554
9. Alfaro-Fernández A, Córdoba-Sellés C, Cebrián MC, et al (2007) First report of tomato torrado virus in tomato in the Canary Islands, Spain. Plant Dis 91:1060
10. Verdin E, Gognalons P, Wipf-Scheibel C, et al (2009) First report of Tomato torrado virus in tomato crops in France. Plant Dis 93:1352
11. Herrera-Vasquez JA, Alfaro-Fernández A, Cordoba-Selles MC, et al (2009) First report of Tomato torrado virus infecting tomato in single and mixed infections with Cucumber mosaic virus in Panama. Plant Dis 93:198
12. Davino S, Bivona L, Iacono G, Davino M (2010) First report of Tomato torrado virus infecting tomato in Italy. Plant Dis 94:1172

13. Gambley CF, Thomas JE, Persley DM, Hall BH (2010) First report of Tomato torrado virus on tomato from Australia. Plant Dis 94:486
14. Verbeek M, Dullemans AM (2012) First Report of Tomato torrado virus infecting tomato in Colombia. Plant Dis 96:592
15. Afechtal M (2020) First report of tomato torrado virus (ToTV) infecting tomato in Morocco. J Plant Pathol 102:1327
16. Moodley V, Gubba A, Mafongoya PL (2020) Emergence and Full Genome Analysis of Tomato Torrado Virus in South Africa. Viruses 12:1167
17. Vucurovic A, Kutnjak D, Mehle N, et al (2021) Detection of Four New Tomato Viruses in Serbia using Post-Hoc High-Throughput Sequencing Analysis of Samples from a Large-Scale Field Survey. Plant Dis
18. Wieczorek P, Budziszewska M, Frackowiak P, Obrepalska-Steplowska A (2020) Development of a New Tomato Torrado Virus-Based Vector Tagged with GFP for Monitoring Virus Movement in Plants. Viruses 12:1195
19. Wieczorek P, Budziszewska M, Obr\kepalska-St\keplowska A (2015) Construction of infectious clones of tomato torrado virus and their delivery by agroinfiltration. Arch Virol 160:517–521
20. Budziszewska M, Wieczorek P, Zhang Y, et al (2014) Genetic variability within the polish tomato torrado virus Kra isolate caused by deletions in the 3'-untranslated region of genomic RNA1. Virus Res 185:47–52
21. Marta B, Henryk P, Aleksandra O-S (2016) Genome Characteristics, Phylogeny and Varying Host Specificity of P olish K ra and R os Isolates of T omato torrado virus. J Phytopathol 164:281–285
22. Roossinck MJ (1997) Mechanisms of plant virus evolution. Annu Rev Phytopathol 35:191–209
23. Hall T, Biosciences I, Carlsbad C (2011) BioEdit: an important software for molecular biology. GERF Bull Biosci 2:60–61
24. Kumar S, Stecher G, Li M, et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549

Figures

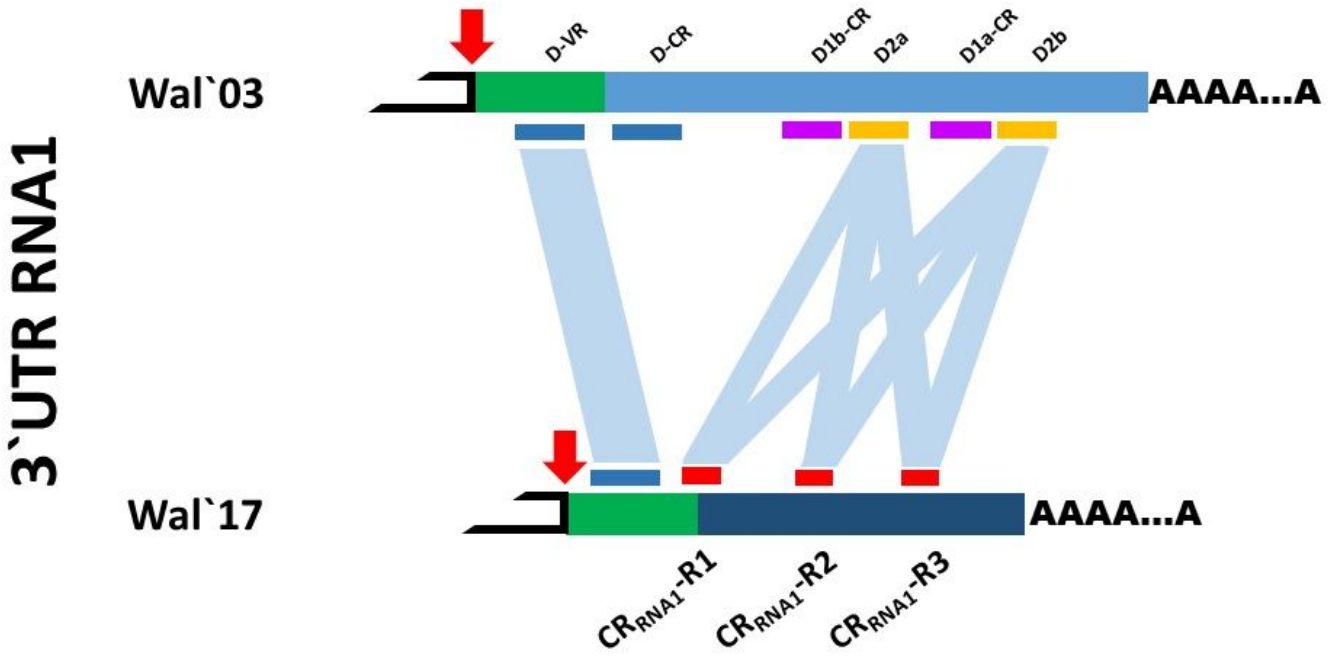

Figure 1

Schematic comparison of 3'untranslated regions of RNA1 of ToTV-Wal'03 and ToTV-Wal'17. The repetitive motifs were indicated with blue, violet, orange (ToTV-Wal'03), and red (ToTV-Wal'03) bars. The red arrow marks the stop codon. The D-VR D-CR regions (the blue bars) are present together only in the genome of ToTV-Wal'03. In RNA1 3'UTR of ToTV-Wal'17 only the D-VR region was identified.

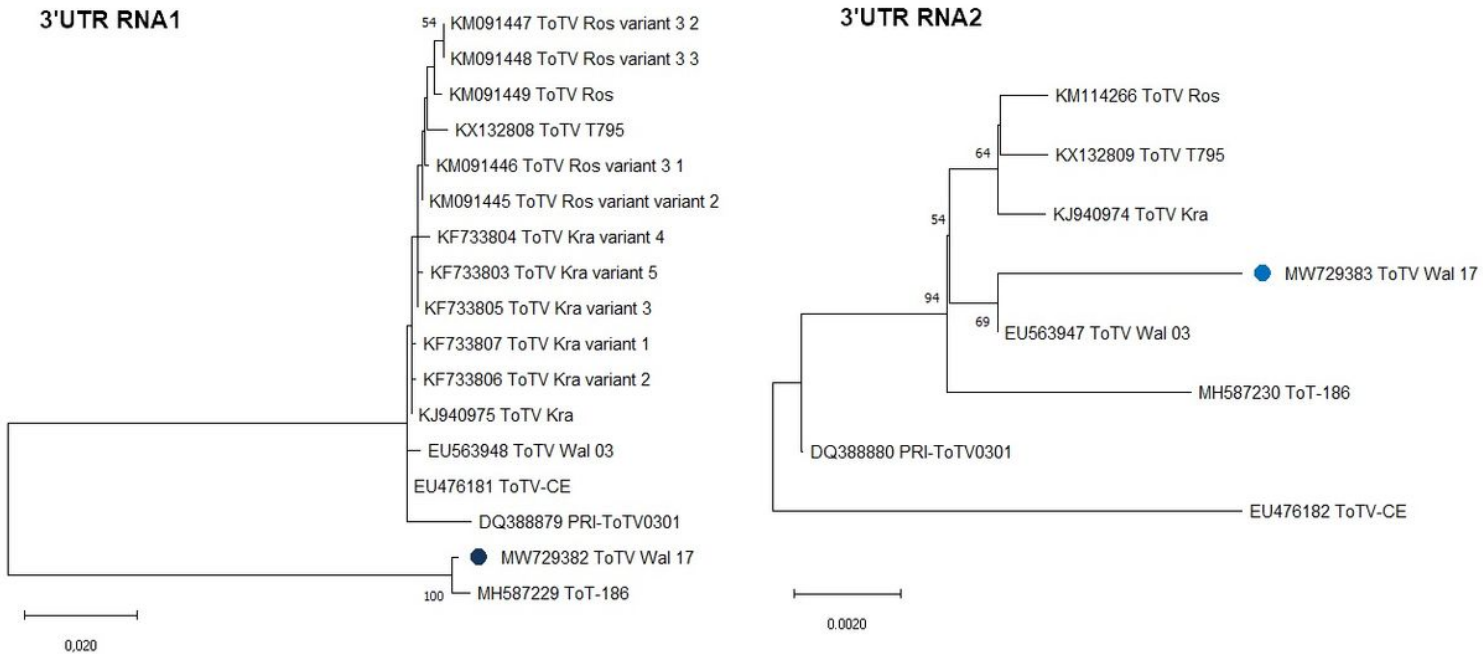

Figure 2

The phylogenetic analyses of 3' untranslated region of RNA1 and RNA2 of ToTV Wal'17. The dendrograms were generated and visualised in MEGA X software, using the Maximum-Likelihood method and 1000 bootstrap value. The scale bars represent a genetic distance. A blue dot indicates the Wal'17 isolate. The accession numbers of sequences deposited in the GenBank NCBI used in analyses are indicated in the trees.