

Tomato chocolàte virus: a new plant virus infecting tomato and a proposed member of the genus *Torradovirus*

Martin Verbeek · Annette Dullemans ·
Hans van den Heuvel · Paul Maris ·
René van der Vlugt

Received: 7 January 2010 / Accepted: 19 February 2010 / Published online: 13 March 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract A new virus was isolated from a tomato plant from Guatemala showing necrotic spots on the bases of the leaves and chocolate-brown patches on the fruits. Structural and molecular analysis showed the virus to be clearly related to but distinct from the recently described *Tomato torrado virus* (ToTV) and *Tomato marchitez virus* (ToMarV), both members of the genus *Torradovirus*. The name tomato chocolàte virus is proposed for this new torradovirus.

In 2007 tomato plants (*Solanum lycopersicum* L.) showing basal leaf necrosis and chocolate-brown spots on the fruits were sampled in the vicinity of Guatemala City, Guatemala. Locally the symptoms on the tomatoes were referred to as ‘Chocolàte’. Electron microscopy on leaf material with typical symptoms showed the presence of spherical virus particles of 28–30 nm in diameter. The virus was mechanically transmissible to tomato cv. ‘Moneymaker’,

Physalis floridana and several *Nicotiana* species including *N. hesperis* ‘67A’ and *N. benthamiana*. Following a slightly modified purification protocol described earlier for *Tomato torrado virus* (ToTV) [10] the virus was purified from *N. hesperis* ‘67A’ for further analysis.

Purified virus was inoculated to tomato plants cv. ‘Moneymaker’ in which it induced symptoms identical to those initially observed in tomato fields in Guatemala. The virus could be isolated from those back-inoculated plants, thus fulfilling Koch’s postulates.

The isolate was tentatively designated tomato chocolàte virus isolate G01 (ToChV-G01) and deposited in the DSMZ Safe Deposit in Braunschweig, Germany, under accession number DSM 22139.

When purified RNA preparations of ToChV were analysed on a denaturing RNA gel [10], two RNA molecules with estimated sizes of 8 kb (RNA1) and 6 kb (RNA2) were observed (results not shown). Subjecting the purified virions to a denaturing polyacryl amide gel electrophoresis we observed three coat proteins of approximately 35, 26 and 24 kDa (Vp35, Vp26 and Vp24, respectively). Based on biological and structural data like symptomatology, particle size and morphology, number and molecular mass of coat proteins, and number of RNAs obtained for ToChV, a possible relationship with viruses from the genus *Torradovirus* (ToTV and *Tomato marchitez virus*; ToMarV [11]) was suspected. Therefore, different upstream primers were derived from ToTV (DQ388879 and DQ388880) as well as ToMarV (EF681764, EF681765, EF063641 and EF063642) sequences and used in combination with a general oligo(dT) primer in a one-tube RT-PCR procedure (Superscript III One-Step RT-PCR system with platinum Taq High Fidelity, Invitrogen) on RNA isolated from purified virus particles. This resulted in a limited number of PCR products. Sequence analyses of the amplicons indeed

M. Verbeek and A. Dullemans contributed to this research equally.

Nucleotide sequence data reported are available in the GenBank database under accession numbers FJ560489 (RNA1) and FJ560490 (RNA2).

M. Verbeek (✉) · A. Dullemans · R. van der Vlugt
Plant Research International, Part of Wageningen UR
(University and Research Centre), Wageningen, The Netherlands
e-mail: martin.verbeek@wur.nl

H. van den Heuvel · P. Maris
De Ruiter Seeds, Bergschenhoek, The Netherlands

Present Address:
H. van den Heuvel
Nickerson-Zwaan, Mado, The Netherlands

revealed low levels of identity with the ToTV and ToMarV RNAs. Based on the obtained sequence information, new ToChV-specific cDNA primers were generated and used to obtain additional sequence information in a 5'-RACE sequence-walking strategy. ToChV-specific primers were used for RT-PCR to confirm sequences of the two RNAs in both orientations.

Nucleotide and amino acid sequence data were analysed and assembled using the DNASTAR package V8 (Lasergene). Sequence comparisons with other viruses were performed with programs from the PHYLIP package. Multiple alignments and phylogenies were performed with the CLUSTALX program after bootstrapping in 1,000 replicates. Neighbour-joining consensus phylogenies were viewed using the NJplot program [8] and printed using TreeView [6].

RNA1 [7,474 nucleotides (nts), excluding the poly(A) tail] contains one open reading frame (RNA1-ORF1) of 6,465 nts encoding a predicted polyprotein of 2,155 amino acids (aa) with a molecular mass of 237 kDa. The 5'- and 3'-untranslated regions (UTR) of RNA1 are 135 and 871 nts in length, respectively. RNA2 of ToChV [5,695 nts, excluding the poly(A) tail] contains two open reading frames (RNA2-ORF1 and RNA2-ORF2) encoding predicted polyproteins of 189 and 1192 aa with molecular masses of 21 and 131 kDa, respectively. The 5'- and 3'-UTRs of RNA2 are 181 and 1,406 nts in length, respectively. The genomic organization is analogous to ToTV and ToMarV [10, 11] with helicase, protease and RNA-dependent RNA polymerase (RdRp) located on RNA1, and a putative movement protein and three coat proteins (Vp35, Vp26 and Vp24) located on RNA2 (Fig. 1).

While writing this report the complete sequence of a virus designated tomato chocolate spot virus (although no acronym has been proposed for this name, we will use ToChSV for the purpose of this paper) became available through GenBank. Like ToChV, ToChSV was reported to be isolated from tomatoes grown in Guatemala. This virus has two ssRNA molecules of 7,492 nts (RNA1:

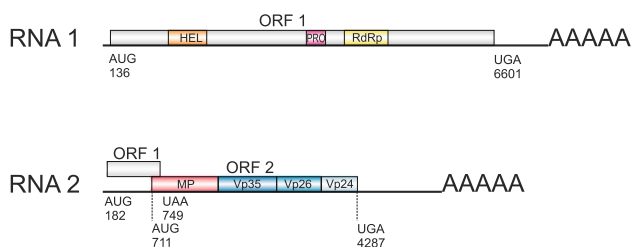
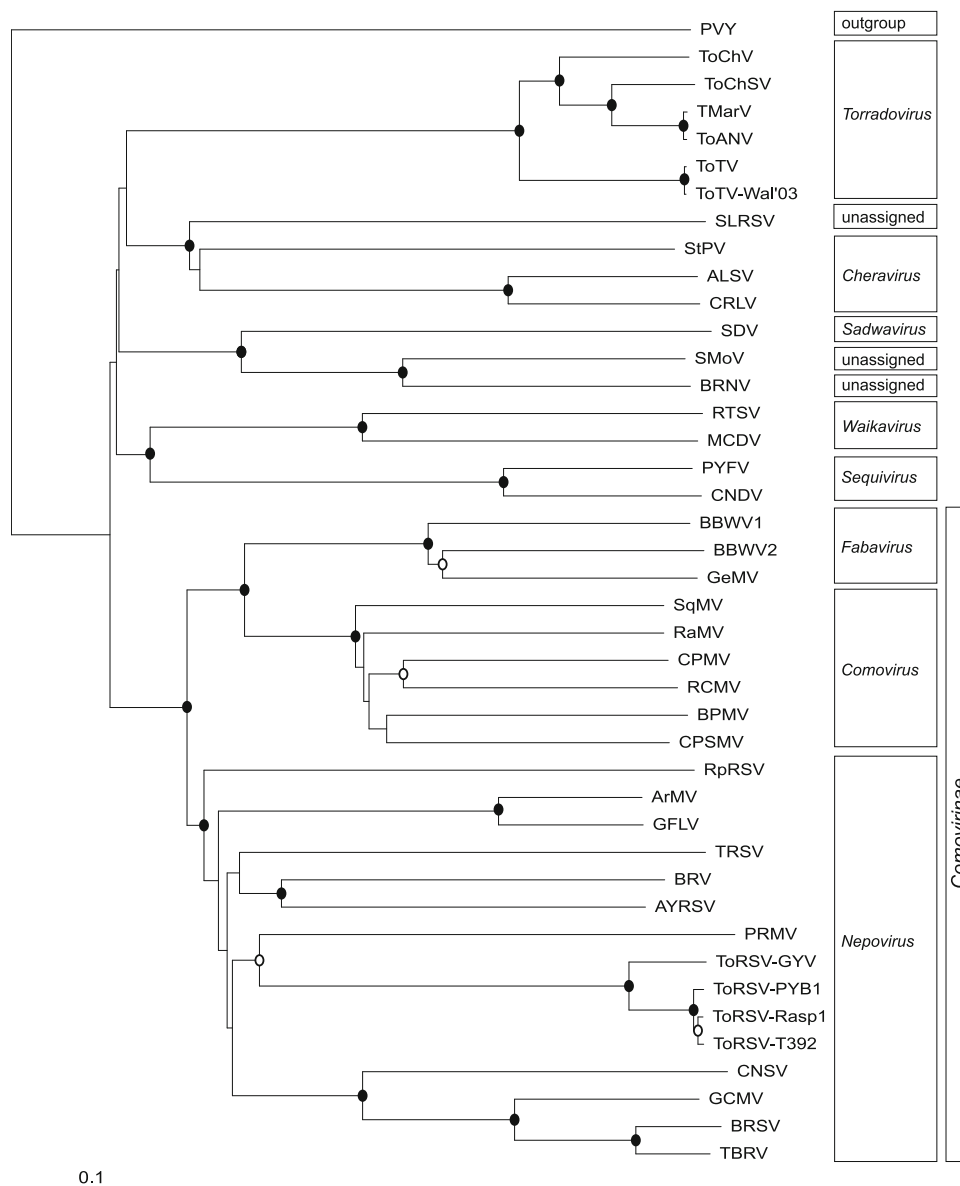


Fig. 1 Genome organization of tomato chocolate virus (ToChV). Positions of start and stop codons are indicated, as well as the relative positions of regions containing motifs of helicase (*HEL*), protease (*PRO*), and RNA-dependent RNA polymerase (*RdRp*) on RNA1 and of the three coat proteins (*Vp35*, *Vp26* and *Vp24*) and putative movement protein (*MP*) on RNA2

Fig. 2 Phylogenetic analysis of tomato chocolate virus (ToChV) and members of the family *Secoviridae* based on the alignment of the region between the protease CG motif and GDD RdRp motif (aa 1,045–1,502 of RNA1-ORF1). Circles indicate nodes supported by bootstrap values above 80% (closed circles) or 60% (open circles); nodes without circles are not supported to these levels. The bar represents a *p*-distance of 0.1. The genera are delineated on the right. Sequences included in the analysis are those of (with virus acronyms, genus and accession numbers in parentheses): apple latent spherical virus (ALSV; *Cheravirus*; AB030940), arabis mosaic virus (ArMV; *Nepovirus*; AY303786), artichoke yellow ringspot virus (AYRSV; *Nepovirus*; AM087671), bean pod mottle virus (BPMV; *Comovirus*; U70866), beet ringspot virus (BRSV; *Nepovirus*; D00322), black raspberry necrosis virus (BRNV; unassigned species in the family *Secoviridae*; DQ344639), blackcurrant reversion virus (BRV; *Nepovirus*; AF368272), broad bean wilt virus 1 (BBWV1; *Fabavirus*; AB084450), broad bean wilt virus 2 (BBWV2; *Fabavirus*; AF225953), carrot necrotic dieback virus (CNDV; *Sequivirus*; EU980442), cherry rasp leaf virus (CRLV; *Cheravirus*; AJ621357), cowpea mosaic virus (CPMV; *Comovirus*; X00206), cowpea severe mosaic virus (CPSMV; *Comovirus*; M83830), cymas necrosis stunt virus (CNSV; *Nepovirus*; AB073147), gentian mosaic virus (GeMV; *Fabavirus*; AB084452), grapevine chrome mosaic virus (GCMV; *Nepovirus*; X15346), grapevine fanleaf virus (GFLV; *Nepovirus*; D00915), maize chlorotic dwarf virus (MCDV; *Waikavirus*; U67839), parsnip yellow fleck virus (PYFV; *Sequivirus*; D14066), peach rosette mosaic virus (PRMV; *Nepovirus*; AF016626), radish mosaic virus (RaMV; *Comovirus*; AB295643), raspberry ringspot virus (RpRSV; *Nepovirus*; AY303787), red clover mottle virus (RCMV; *Comovirus*; X64886), rice tungro spherical virus (RTSV; *Waikavirus*; M95497), satsuma dwarf virus (SDV; *Sadwavirus*; AB009958), squash mosaic virus (SqMV; *Comovirus*; AB054688), stocky prune virus (StPV; *Cheravirus*; DQ143874), strawberry latent ringspot virus (SLRSV; unassigned species in the family *Secoviridae*; AY860978), strawberry mottle virus (SMoV; unassigned species in the family *Secoviridae*; AJ311875), tobacco ringspot virus (TRSV; *Nepovirus*; U50869), tomato black ring virus (TBRV; *Nepovirus*; AY157993), tomato ringspot virus (ToRSV; *Nepovirus*; AF135410 (isolate T392), AF135409 (isolate Rasp1), AF 135408 (isolate PYB1) and AF135407 (isolate GYV), tomato torrado virus [ToTV; *Torradovirus*; DQ388879 (isolate PRI-ToTV0301) and EU563948 (ToTV-Wal'03)], tomato marchitez virus [ToMarV; *Torradovirus*; EF681764 (isolate PRI-TMarV0601) and EF063641 (ToANV)], tomato chocolate spot virus (ToChSV; *Torradovirus*; GQ305131). The sequence of Potato virus Y (PVY; *Potyvirus*; X12456) was used as an outgroup in the analyses

GQ305131) and 5,109 nts (RNA2: GQ305132) with a similar genome organization as the other three torradoviruses including the typical ORF1 on RNA2.

In previous studies [10, 11] phylogenetic comparisons based on the aa region between the CG protease motif [2] and the GDD RdRp active site [1] in the RNA1-ORF1 were used to determine the relationships between torradoviruses and other viruses of the family *Secoviridae* (see for current taxonomy: <http://www.ictvonline.org/virusTaxonomy.asp?>). This region is proposed to be a good taxonomic predictor for classifying picorna-like viruses [4, 5]. A similar analysis was performed to determine the taxonomic position of ToChV and ToChSV. The resulting dendrogram (Fig. 2) shows that both ToChV and ToChSV cluster together with ToTV and ToMarV in the genus *Torradovirus*.



Sanfaçon et al. [7] defined preliminary criteria for defining species within the genus *Torradovirus*. Torradoviruses are distinct species if they share less than 80% aa identity in the 3C-like Proteinase-RNA-dependent RNA polymerase (Prot-Pol) region of RNA1-ORF1 or less than 75% aa identity in the CP region of RNA2-ORF2.

Table 1 lists the percentages of identity between ToChV and other isolates of torradoviruses for the ORFs and regions of RNA1 and RNA2. Strictly based on the criterion set for the Prot-Pol region of RNA1 ToChV and ToChSV could be considered isolates of ToMarV. Comparisons of the complete CP sequences between the viruses also suggest that ToChV, ToMarV and ToChSV groups as one species as they share between 82.0 and 87.1% aa identity. It should be noted, however, that also ToTV and ToMarV

could be considered one species as they share 75.3% identity in this region. The criteria as set by Sanfaçon et al. [7] are based on only a limited number of full sequences (2 isolates of ToTV [3, 10] and one of ToMarV [11]) and an incomplete sequence of an isolate of ToMarV (tomato apex necrosis virus; ToANV [9]). As the authors indicate these preliminary criteria may need to be adjusted as more viruses become characterized.

To obtain a better understanding of the possible taxonomic relations between the four torradovirus sequences we performed more elaborate comparisons of the different coding and non-coding regions that have now become available. Overall levels of nt sequence identity between all viruses are well below 80% and mostly around or below 70% while for isolates of the same virus identity levels are

Table 1 Comparison of tomato chocolate virus sequences with other members of the genus *Torradorovirus*

| Virus | Isolate | RNA1 | | | | | RNA2 | | | | | | | | | |
|-----------------------------|----------------|------------------|------------|----------|-----------------|-----------------|------------------|------|-----------|------|------|------|-----------------|-----------------|------------|-----------|
| | | Overall (length) | ORF1 | | UTR | | Overall (length) | ORF2 | | UTR | | | | | | |
| | | | Entire ORF | Prot-Pol | 5'-UTR (length) | 3'-UTR (length) | | MP | CP region | Vp35 | Vp26 | Vp24 | 5'-UTR (length) | 3'-UTR (length) | | |
| Tomato chocolate virus | ToChV-G01 | - (7,474) | - | - | - (135) | - (871) | - (5,695) | - | - | - | - | - | - | - | - (181) | - (1,406) |
| Tomato torrado virus | PRI-ToTV0301 | 63 (7,808) | 64 | 79 | 76 (106) | 54 (1,210) | 62 (5,403) | 60 | 74 | 67 | 83 | 74 | 74 | 55 (181) | 45 (1,091) | |
| Tomato torrado virus | ToTV-Wal'03 | 63 (7,829) | 64 | 79 | 76 (107) | 54 (1,229) | 62 (5,404) | 61 | 74 | 68 | 83 | 74 | 74 | 55 (181) | 44 (1,092) | |
| Tomato marchitez virus | PRI-TMarV 0601 | 69 (7,229) | 74 | 85 | 61 (140) | 74 (625) | 69 (4,906) | 64 | 83 | 80 | 87 | 83 | 83 | 60 (138) | 60 (652) | |
| Tomato marchitez virus | ToANV | 69 ? (6,478) | 74 | 85 | ? (?) | 73 (627) | 71 ? (3,013) | ? | 82 | 78 | 88 | 83 | 83 | ? (?) | 60 (647) | |
| Tomato chocolate spot virus | - | 71 (7,492) | 73 | 82 | 64 (138) | 96 (867) | 70 (5,109) | 66 | 83 | 79 | 87 | 85 | 85 | 57 (171) | 78 (814) | |

Identities (%) for the amino acid sequences are given in bold and for the nucleotide sequences in roman, lengths (nts) in brackets. The question mark (?) indicates lack of sequence information on the tomato apex necrosis isolate of ToMarV

significantly higher (ToTV and ToTV-Wal-'03: 99.0% for RNA1 and 99.1% for RNA2; ToMarV and ToANV: 95.3% for RNA1 and 79.9% for RNA2). However, since for ToANV only partial sequences are published (the 5'-ends of both RNAs are missing), these numbers may change if additional sequence information on ToANV becomes available. Overall sequence identity between ToChV and the other torradoviruses ranges from 62 to 71% (Table 1).

The 5'- and 3'-UTRs clearly differ between the four (tentative) species of the genus *Torradorovirus*. The 5'-UTRs of both RNA1 and RNA2 of ToChV and ToChSV are fairly similar in length (135 vs. 138 for RNA1 and 181 vs. 171 nts for RNA2) but overall sequence identities are only 64 and 57% for RNA1 and RNA2, respectively (Table 1). The lengths of the 5'-UTRs of the RNAs1 and RNAs2 of ToTV and ToMarV are clearly different from ToChV with levels of overall identity varying between 44 and 76% (see Table 1).

Direct sequence comparisons between the total 3'-UTRs of RNA1 and RNA2 of the four torradovirus sequences are difficult because of large differences in their 3'-UTR lengths (Table 1). Identity levels in the 3'-UTR of RNA1 and 2 are generally lower than for the 5'-UTR and again lower in RNA2 in comparison to RNA1 (results not shown).

Overall sequence identity levels to the 3'-UTR of ToChV vary between 54% (ToTV) and 96% (ToChSV) for RNA1 and 44% (ToTV-Wal'03) and 78% (ToChSV) for RNA2 (Table 1). Remarkably in both 3'-UTRs identity levels for isolates of one virus are significantly higher, i.e. above 99% for ToTV and ToTV-Wal'03 and around 89% for ToMarV and the tomato apex necrosis isolate of ToMarV.

More detailed analyses of both 3'-UTRs of each torradovirus show that they share a unique species specific region (>99%) that is highly conserved between the 3'-terminal ends [directly preceding the poly(A) tail] of the two RNAs. Between species this region is clearly different in length but nearly identical between isolates of one virus; ToChV 787 nts, ToTV 989 nts, ToTV-Wal'03 990 nts, ToMarV 552 nts and ToANV 554 nts. Remarkably, ToChSV was the only virus lacking this conserved region at the end of the 3'-UTR of both RNAs. There is no obvious explanation for this but it should be noted that the 772 nts directly preceding the poly(A) tail of the 3'-UTR of RNA1 of this virus share 99.2% sequence identity with the last 772 nts of both RNA1 and RNA2 of ToChV. This suggests that the sequence of the 3'-UTR of RNA1 of ToChSV might be directly derived from ToChV.

With respect to the coding regions, we made comparisons of the RNA1-ORF1 (entire ORF, helicase region and Prot-Pol region), the RNA2-ORF1 and RNA2-ORF2 (MP and CP region) of ToChV to the other torradoviruses

(Table 1). RNA1-ORF1 of ToChV shows aa sequence identities of 64% (ToTV) to 74% (ToMarV). The aa sequence identity for the helicase region, which is a highly conserved region, is much higher (95–96%, not shown in Table 1). The aa sequence identity in the Prot-Pol region varies between 79% (ToTV) and 85% (ToMarV).

RNA2-ORF1 of ToChV shows aa identities of 60% (ToTV) to 66% (ToChSV). In the second ORF of RNA2 the aa sequence identities in the MP region vary between 65% (ToTV) and 75% (ToChSV) and in the CP region between 74% (ToTV) and 83% (ToMarV and ToChSV). The different predicted CPs of ToChV share aa sequence identities of 67–88% with those of ToTV and ToMarV. Remarkably, the identity between the Vp26 analogues of the different viruses is significantly higher (over 83%) than for the other two CPs (67–80% and 74–85% for the Vp35 and Vp24, respectively) (Table 1).

In conclusion, there are very distinct sequence differences between the four (proposed) torradovirus species (ToTV, ToMarV, ToChV and ToChSV). However, isolates of one virus species (as seen in the two isolates of ToTV and ToMarV) have high sequence similarities in the coding and non-coding regions, in which the (tentative) torradovirus species differ significantly. The RNA2-ORF1 is a region very similar in length between the currently known torradoviruses and unique for the genus *Torradovirus*. It displays a significant sequence variation between the four torradovirus sequences with levels of identities ranging from 60.4 to 73.9% and is nearly identical for the two isolates of ToTV (99.5%).

The 3'-UTRs of the torradoviruses clearly differ in length and identity between species but harbour a unique, nearly identical region in the two RNAs of one species. This virus-specific region differs significantly between different species both in length and in identity (between species <65% and for isolates of the same species from 85 to 99%).

Based on our observation we propose that identity levels below 75% in the RNA2-ORF1 and 85% in the conserved region in the 3'-UTR of both RNAs of each torradovirus can serve as additional to the current tentative taxonomic criteria [7] which may need to be updated now that additional information has become available.

The analysis presented above and application of the newly proposed additional demarcation criteria leads us to

conclude that tomato chocolate virus is a new species in the genus *Torradovirus*.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Argos P (1988) A sequence motif in many polymerases. *Nucleic Acids Res* 16:9909–9916
- Bazan JF, Fletterick RJ (1988) Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. *Proc Natl Acad Sci USA* 85:7872–7876
- Budziszewska M, Obrepalska-Stepłowska A, Wieczorek P, Pospieszny H (2008) The nucleotide sequence of a Polish isolate of Tomato torrado virus. *Virus Genes* 37:400–406
- Le Gall O, Sanfaçon H, Ikegami M, Iwanami T, Jones T, Karasev A, Lehto K, Wellink J, Wetzel T, Yoshikawa N (2007) Chera-virus and Sadwavirus: two unassigned genera of plant positive-sense single-stranded RNA viruses formerly considered atypical members of the genus Nepovirus (family Comoviridae). *Arch Virol* 152:1767–1774
- Le Gall O, Christian P, Fauquet CM, King AMQ, Knowles NJ, Nakashima N, Stanway G, Gorbalenya AE (2008) Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Arch Virol* 153:715–727
- Page RDM (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12:357–358
- 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 genera Chera-virus and Sadwavirus, and the proposed genus Torradovirus. *Arch Virol* 154:899–907
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882
- 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 Dis* 91:932–941
- Verbeek M, Dulleman AM, Van den Heuvel JFJM, Maris PC, Van der Vlugt RAA (2007) Identification and characterisation of tomato torrado virus, a new plant picorna-like virus from tomato. *Arch Virol* 152:881–890
- Verbeek M, Dulleman AM, Van den Heuvel JFJM, Maris PC, Van der Vlugt RAA (2008) Tomato marchitez virus, a new plant picorna-like virus from tomato related to tomato torrado virus. *Arch Virol* 153:127–134